Application of the Non-Target Screenings using LC-ESI-HRMS in Water Analysis

(Edition 2.0 2025)

Water Chemistry Society

Division of the German Chemical Society

ISBN: 978-3-947197-24-8

Imprint

Edition 2.0 2025

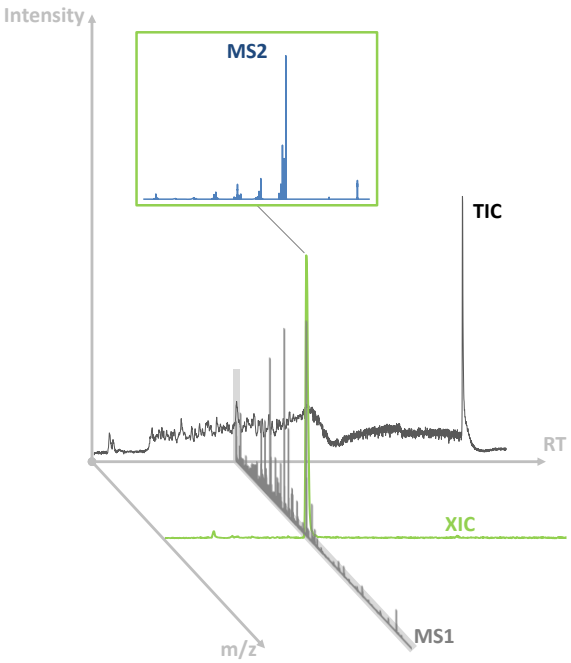
Published in Koblenz, October 2025

Responsible for the Content:
Dr. Wolfgang Schulz
Aalen University - Technology, Business and Health
Beethovenstrasse 1
73430 Aalen
Wolfgang.Schulz@hs-aalen.de

Publisher:
Expert Committee "Non-Target Screening" of the
Water Chemistry Society
Federal Institute of Hydrology
Am Mainzer Tor 1
56068 Koblenz
T +49 69 7917-368
E: wasserchemische-gesellschaft@go.gdch.de
Web: https://www.gdch.de/netzwerk-strukturen/fachstrukturen/wasserchemische-gesellschaft.html

Guide

Application of Non-Target Screening by Means of LC-ESI-HRMS in the Water Analysis



Edition 2.0 2025

GDCh

WASSERCHEMISCHE GESELLSCHAFT
WG

Water Chemisty Society
Expert Committee 'Non-Target Screening'

Non-Target Screening in Water Analysis

Guide to the use of LC-ESI-HRMS for screening examinations

Edition 2.0 2025

This guideline was developed by the members of the Technical Committee 'Non-Target Screening' of the Water Chemistry Society.

Members of the expert committee

Head: Bader, Tobias	**Zweckverband Landeswasserversorgung**
Jewell, Kevin	**Federal Institute of Hydrology**
Armbruster, Dominic	DVGW Water Technology Center
Brüggen, Susanne	Landesamt für Natur, Umwelt und Klima Nordrhein-Westfalen
Cunha, Ricardo	Institute for Energy and Environmental Technology e.V. (IUTA)
Götz, Sven	Hessenwasser GmbH & Co KG
Härtel, Christoph	Ruhrverband
Käberich, Merle	DVGW Water Technology Center
Kronsbein, Anna Lena	Federal Environment Agency
Kunkel, Uwe	Bavarian Environment Agency
Letzel, Thomas	Analytical Research Institute for Non-Target Screening GmbH (AFIN-TS)
Liebmann, Diana	Laboratory of Berliner Wasserbetriebe (BWB)
Liesener, André	Westfälische Wasser- und Umweltanalytik GmbH
Logemann, Jörn	Institute for Hygiene and Environment , Hamburg
Lucke, Thomas	Zweckverband Landeswasserversorgung
Macherius, André	Bavarian Environment Agency
Merkus, Valentina	Hamburger Wasserwerke GmbH
Petri, Michael	Zweckverband Bodensee-Wasserversorgung
Reineke, Anna	Westfälische Wasser- und Umweltanalytik GmbH
Renner, Gerrit	University of Duisburg-Essen
Ruppe, Steffen	Amt für Umwelt und Energie Basel
Scheurer, Marco	LUBW Landesanstalt für Umwelt Baden-Württemberg
Schlüsener, Michael	Federal Institute of Hydrology
Schulz, Wolfgang	Aalen University of Applied Sciences
Singer, Heinz	Eawag-Swiss Federal Institute of Aquatic Science and Technology
Türk, Jochen	Cooperation laboratory of Ruhrverband, Emschergenossenschaft and Lippeverband
Zahn, Daniel	Helmholtz Center for Environmental Research
Zwiener, Christian	University of Tübingen

Citation recommendation

Citation of the guideline as follows: "W. Schulz *et al.*, Guide: Application of Non-Target Screening by Means of LC-ESI-HRMS in the Water Analysis (2025). ISBN: 978-3-947197-24-8, Download at https://en.gdch.de/network-structures/gdch-structures/society-of-water-chemistry/publications/wchg-committees.html"

Table of contents

List of illustrations

List of tables

1 Introduction

The use of liquid chromatography (LC) coupled with high-resolution mass spectrometry (HRMS) enables the qualitative detection and quantitative determination of organic trace substances [1-5]. A distinction is made between quantitative target analysis and qualitative non-target screening (NTS). In target analysis, previously defined substances to be detected in a (water) sample are analyzed, and their concentrations are determined with reference substances. Both known substances (suspect-target screening) and previously undetected and often unknown substances can be detected with NTS. The retrospective evaluation of, for example, newly discovered or previously unconsidered organic trace substances is a particular advantage of HRMS compared to using low-resolution mass spectrometers [4].

This guide is intended to make it easier for new users to get started with water analysis using LC-HRMS and to serve as an orientation aid. The authors' numerous practical experiences have been incorporated into the individual chapters. The prerequisites and requirements for the measurement technology, evaluation, and interpretation of the data are described firsthand. The document was written by users for users, points out pitfalls and the various influences on the result of the NTS, and provides information on possible solutions. The guideline describes many generally valid approaches from sampling to the evaluation of NTS results but focuses on the implementation of quality-assured, practice-oriented analysis of water samples. This sets it apart from the NORMAN network guide [6], which also deals with other sample matrices and focuses on NTS's technical and instrumental aspects.

Compared to the first edition of the guideline from 2019, this second edition additionally handles validation and quality assurance in a separate chapter. In addition, the chapter on statistical analysis has been significantly expanded. Specific suggestions for quality assurance are presented here, which should be applied to all typical quantitative and qualitative issues in the analysis of drinking water, groundwater, surface water, and wastewater samples (Table 1.1)

Table 1.1: Overview of typical tasks in water analysis[1]

Target analytics	Suspect target screening	Non-target screening
• Monitoring of organic trace substances to control limit values • Monitoring of organic trace substances to determine trends • Monitoring organic trace substances after contamination (accident, fire, etc.) • Monitoring of individual process steps in wastewater and drinking water treatment (e.g., breakthrough of an adsorption filter, elimination performance of individual process stages)	• Search for expected substances without reference substances (e.g., industrial chemicals, transformation products, etc.) • Search for substances with specific structural properties (elements in the molecule such as S, Cl, Br or functional groups such as -COOH) • Comparison of positive findings from tests by other laboratories or from literature data • Retrospective consideration of archived HRMS measurements for information on new substances - Quick assessment of the occurrence of the component at the measuring point under consideration - Decision-making basis for the expansion of monitoring programs	• Search for and characterization of additional substances present (beyond monitoring) • Determination of differences (concerning organic trace substances) between several samples (hydrogeology, temporal course, process consideration concerning removal or formation of unknown substances); application of statistical • Description of processes concerning the behavior of organic trace substances • Detection and characterization of transformation products (e.g., known starting compounds) • Detection/occurrence of substances as a result of an event (determination of cause, e.g., toxicity - fish mortality, odor - taste; heavy rain, accident, fire, etc.) • Expansion/adaptation of monitoring programs (dynamic monitoring) • Identification of unknown substances with the aid of further information (database comparison, comparison of MS/MS spectra from literature data or in-silico *fragmentations*) and measurements (reference substances, use of orthogonal techniques such as NMR or Raman spectroscopy) • Assignment of samples to a specific group (class), e.g., origin

[1]Revised from "Possibilities of high-resolution mass spectrometry (HRMS), use of suspect and non-target analysis in the monitoring practice of raw and drinking water" DVGW-Information Water No. 93

2 Area of application

This guideline is intended to highlight fundamental aspects of applying high-performance liquid chromatography coupled with high-resolution mass spectrometry. In addition to the technical aspects of the equipment and possible contamination during sampling and measurement, this also includes data evaluation, validation, and quality assurance measures. The guide is intended to assist the user in developing methods and interpreting the results [7].

3 Terms and abbreviations

The most important terms of mass spectrometry and high-performance liquid chromatography, along with their definitions, are summarized in Table 3.1.

Table 3.1: Compilation of abbreviations and terms used in mass spectrometry and high-performance chromatography [8]

AIF, MSMS^{all}, MS^{E}	*All Ion Fragmentation*
Accurate mass	An ion's accurate mass is determined experimentally using a high-resolution mass spectrometer (and, if necessary, recalibrated using a reference mass standard).
APCI	*Atmospheric pressure chemical ionization*
DDA **IDA**	*Data-Dependent Acquisition,* *Information-Dependent* data-dependent acquisition A predetermined number of precursor ions, whose *m/z values* were recorded in an overview scan, are selected using predefined rules and subjected to a second stage of mass selection in an MS/MS analysis [8].
DIA	*Data-Independent Acquisition* data-independent acquisition In each analysis cycle, all precursor ions that reach the mass spectrometer are fragmented [9].
ESI	*Electrospray ionization*
Exact mass	The exact mass of an ion or molecule is the calculated mass, taking into account the given isotopic composition.
Feature	Features are peak-shaped signals that are defined by their accurate mass (*m/z*) and retention time (RT) and fulfill the selected criteria for peak finding (e.g., intensity threshold).
FT-ICR-MS	*Fourier transform ion cyclotron resonance mass spectrometer*
Full scan (FS)	Recording of a mass spectrum over the entire mass range with positive, negative, or alternating ionization (*switch*)
Targeted MS^2 **ddMS^2**	*Dedicated* fragmentation of individual ions to obtain the purest possible fragment ion spectra

	Ions for fragmentation can be selected from a predefined list or automatically based on a survey scan according to defined criteria.
HILIC	*Hydrophilic Interaction Liquid* MS-capable alternative to normal phase chromatography for the separation of strongly polar compounds consisting of a polar stationary phase (similar to normal phase chromatography; partly in combination with cation/anion exchange functions) using conventional RP eluents (mostly water and acetonitrile)
(HP)LC	***(****High-Performance) Liquid Chromatography*
HRMS	*High-Resolution Mass Spectrometry* High resolution is present if the resolution is > 10,000 according to the valley definition or > 20,000 according to the definition via the peak half-width.
Isotope pattern, isotope	The pattern forms in the mass spectrum is due to the mass spectrometric separation of the different isotopes of the atoms in the molecule. The isotope pattern depends on the frequency of the individual atoms in the molecule and the resolution of the mass spectrometer.
LC-HRMS	*Liquid Chromatography - High-Resolution Mass Spectrometry* Liquid chromatography coupled with high-resolution mass spectrometry
LIMS	*Laboratory information and management system*
m/z	Abbreviation for mass-to-charge ratio Mass number divided by charge number (dimensionless)
Mass defect	The mass defect of an atom, molecule, or ion is the difference between the nominal and monoisotopic mass. Most organic molecules have a positive mass defect, often composed of atoms with almost negligible negative (e.g., O, F) or minor positive mass defects (e.g., H, N). Some elements, such as chlorine and bromine, have significant negative mass defects. For the most common carbon isotope ($^{(12)}C$), the mass defect is zero by definition.
Mixed Mode	LC column material (stationary phase) with a combination of different functionalities to form hydrophobic and ionic (ion exchange) interactions
Monoisotopic mass	The exact mass of an ion or molecule is calculated from the elements' most frequently occurring natural isotopes. In this context, the monoisotopic mass of molecules or ions is also called the exact mass.
MRM	*Multiple Reaction Monitoring*
MS	*Mass spectrum* Two-dimensional representation of the signal intensity (y-axis) plotted against the *m/z* ratio (x-axis)

MS² (MS/MS)	Recording of production spectra (fragment ion spectra) by molecular fragmentation
Nominal mass	The nominal mass of an element is the whole number of the mass of its most common isotope, for example, 12 u for carbon and 35 u for chlorine. To calculate the nominal mass of a molecule or ion, the nominal masses of the elements are multiplied by the number of atoms of each element in the molecule or ion.
NTS	*Non-target screening* Non-targeted analysis method without limitation to specified substances. All substances that can be detected chromatographically and mass spectrometrically with the analytical method are detected.
Precursor ions	Ion that reacts to form specific product ions or suffers certain neutral losses
QA	*Quality assurance*
Resolution **Resolving power** Both terms are used synonymously in practice.	The **mass resolution** is the quotient of the mass m determined in the mass spectrometer and the difference Δm between two neighboring masses at the point m that can still be separated from each other (*m/Δm*). The mass difference Δm can be measured from peak to peak (same intensity at an overlap of 10% valley height) or by the peak width, usually at 50% of the peak height (full width at half maximum, FWHM) [8]. The **resolving power** is the ability of an instrument to distinguish between two peaks that differ by a small *m/z value* and is defined as the peak width *Δm*.
RP	*Reversed-phase* Reversed-phase in high-performance liquid chromatography
SRM	*Selected-Reaction Monitoring*
SWATH-MS	*Sequential Window Acquisition of all Theoretical Mass Spectra*
TOF-MS	*Time of Flight Mass Spectrometer*
u	The *atomic mass unit* is defined as the twelfth part of the mass of a carbon atom in its ground state: 1 u = 1.660 539 040 10^{-27} kg Equivalent to Da (Dalton)
U(H)PLC	*Ultra (High) Performance Liquid Chromatography* High-performance liquid chromatography with very high chromatographic separation performance on columns with small particle sizes (< 2 µm) and at column pressures of up to 1500 bar

4 Basis of the procedure

The method is based on high-performance liquid chromatography coupled with high-resolution mass spectrometry (LC-HRMS). This makes it possible to detect the ions formed in the ion source in the selected mass range at any point in the chromatogram and to determine their accurate mass [1, 10]. Mass detection can be carried out using a time-of-flight mass analyzer (TOF), an Orbitrap, or another high-resolution mass spectrometer (FT-ICR, sector field MS) (see chapter 8.4).

The identification of substances requires the measurement of MS^2 spectra with accurate mass for individually selected (MS/MS or $ddMS^2$) or, if possible, simultaneously for all precursor ions (MS/MS^{all} or AIF or DIA). The evaluation of the data obtained depends on the question and is divided into suspect-target and non-target screening (Table 1.1).

4.1 Non-target screening

LC-HRMS data is searched for features in NTS using suitable peak-finding software. Due to isotope peaks, the formation of different adductions of a molecule in the ion source, and possible in-source fragmentation, componentization is required i.e., a summary of all signals initially originating from one component (see chapter 10.1.5). A blank value correction is also necessary to remove false-positive features (see chapter 10.1.4). Alignment is also required to compare different samples (see chapter 10.1.3). This is usually followed by generating possible summation formulas using the accurate masses of the features and, if these can be detected with sufficient concentration or sensitivity, the corresponding isotopes (see chapter 10.1.6). In addition to mass accuracy, the resolution of the mass spectrometer plays an essential role in reducing the number of possible sum formula proposals. Reference should also be made to the "*Seven Golden Rules"* for reducing molecular formulae to chemically meaningful proposals [11]. Various tools are available to identify and interpret the features. The MS^2 information recorded for the features is essential for determining structures. This can be compared with MS^2-spectra databases (in-house, Massbank [12], mzcloud [13]). In addition, substance databases such as PubChem [14], ChemSpider [15], etc., can also be queried, and possible hits can be further prioritized using in-silico tools.

An example of this is the FOR-IDENT platform, which links the *in-silico* tool Metfrag [16] with the substance database STOFF-IDENT [17] (see chapter 10.2.1.1). The number of possible structure proposals for individual features decreases successively as more information is included in the queries. As it is only possible to clearly identify a feature with a reference standard, the classification of possible formula, structure, and substance proposals into different categories based on the degree of fulfillment of various criteria has proven successful (see chapter 10.2.1). In addition, metadata, statistical methods, and comparing results from different samples (even without identification) can provide significant assistance in answering the analytical question (e.g., prioritizing relevant features).

The focus of the methods described here is on detecting hydrophobic to moderately polar substances that can be separated using RP-LC and ionized using ESI. This, therefore, covers a large proportion of common pharmaceuticals, pesticides, industrial chemicals, and their transformation products. By using HILIC or mixed-mode materials as stationary phase, the range can be extended to (very) polar substances. However, critical substances can range from ionic (very polar) compounds (e.g., TFA) to very hydrophobic and poorly ionizable compounds (e.g., benzene). In addition, critical substances can cover a wide mass range, from, e.g., cyanide (26 Da) to oligopeptides (kDa range). It is impossible to cover all potentially critical compounds

with a single analytical method. Therefore, selecting the method according to the problem at hand is necessary. Separation techniques such as gas chromatography, ion chromatography, capillary electrophoresis, and supercritical fluid chromatography are also increasingly used in connection with NTS.

4.2 Suspect target screening

Suspect-target screening provides a list of relevant substances or substance groups to answer the question. This list can also be generated from corresponding substance databases or via suitable platforms (e.g., FOR-IDENT [18], CompTox Chemicals Dashboard [19]) using task-specific criteria (e.g., drugs, neurotoxic substances). Suitable software is used to search the LC-HRMS data of the sample(s) for indications of the presence of these suspects. Various strategies can be used for this, such as exact masses or the input of sum formulas. Verifying positive indications (identification) generally requires an MS^2 spectrum of the sample and reference substance or corresponding information from the literature.

5 Blank signals

All types of blank value signals must be avoided or minimized. The origin of blank value signals can be assigned to different work steps or sources. The causes of blank value signals and their avoidance in the individual work steps are explained in more detail below.

5.1 Blank values -signals due to sampling

Blank value signals from sampling should be kept to a minimum. Sampling containers should only be classified for sampling drinking, surface, or wastewater to avoid cross-contamination. This avoids using a glass bottle previously filled with wastewater for drinking water sampling. All sample containers or glassware should be baked in a bake-out oven for at least 4 hours or overnight at a minimum of 450 °C to minimize blank values. If possible, inert containers of glass or stainless steel should be used. If this is not possible, e.g., for technical reasons (mixed samples via automatic samplers, temperature resistance), containers made of plasticizer-free plastics or well-rinsed or old, contamination-free plastic containers should be used. It is crucial to generate a representative blank sample for data evaluation, e.g., to fill pure water into the sample container via the automatic sampler.

Any sample handling, such as decanting, pipetting, or enrichment, can lead to contamination with organic trace substances (also by laboratory personnel, e.g., through skin protection or skin care products and cigarette smoke residues).

5.2 Blank signals through the analysis

Open handling (e.g., decanting) should be avoided to avoid contamination. The addition of auxiliary materials (e.g., formic acid to the eluents used solvents to the analysis sample) should ideally be carried out using heated glassware (see chapter 8.1). The devices and analysis systems used should be regularly maintained and checked for possible contamination, e.g., by lubricants or the materials used (e.g., pipes, seals).

5.3 Measurement of blank samples

Regular blank value measurements must be used to ensure the devices and chemicals are in perfect condition. For example, a sample blank value and/or a system blank value can be used to check the blank value. A laboratory pure water sample or synthetic buffered water with identical sampling, storage, transportation, and preparation procedures as the original sample is generally used as the sample blank. The system blank is the chromatogram without sample injection (zero injection). The total ion current chromatograms obtained can be assessed by comparing the signal intensity (see appendix C.3). To assess the blank value, an additional evaluation must be carried out according to 10.1.2. A blank value check must be carried out in each measurement sequence. When measuring samples with unknown contamination, a blank value measurement is recommended between the injections to avoid or detect carryover.

6 Sampling

6.1 General

The taking of water samples is described in various standards for a wide range of parameters and parameter groups [20]. A check for contamination or losses (e.g., due to adsorption or instability of the sample during sample transport to the laboratory) can be carried out based on individual components, but not in the case of NTS for all known and unknown substances in the sample. Therefore, basic precautions must be taken when taking samples [21].

The required sample volume depends on the sample preparation and the injection volume. Stabilization by adding acid or sodium azide (microbiology) can lead to contamination and chemical reactions. Immediate cooling of the sample to approx. 4 °C and analysis as quickly as possible is recommended. If this is not possible, the samples should be frozen at a maximum of -18 °C until analysis. This also applies to retained samples. Possible losses due to freeze/thaw cycles must be taken into account.

6.2 Quality assurance during sampling

Implementing quality assurance measures during sampling can reduce misinterpretations of the measurement results. A suitable quality assurance measure must be tested for the respective task. The use of so-called field blanks has proven itself in some tasks for quality control of sampling, such as pump sampling. The field blank sample is filled on site (e.g., ultrapure water). This allows contamination to be detected during sampling or transportation. In the case of complex sample transportation, an additional transport blank is helpful for each transport container (cool box).

6.3 Sample designation/sample description

The sample designation must be chosen so that all data (raw data, evaluation) can be assigned to the sample without any doubt. Using a unique laboratory number that appears without gaps in all file designations and documents is advisable. Table 6.1 contains examples of accompanying sample information. For further information, please refer to the relevant standards on the different types of sampling [20, 22, 23]

Additional information or special features (meta information) during sampling should be attached to the documents. This facilitates the interpretation of the screening data. For this purpose, the task must be defined and known to the sampler.

Table 6.1: Exemplary compilation of sample accompanying information

Information	**Description/Example**
Sampling point	Exact designation: For example, river kilometers, name of the groundwater measuring point, geographical coordinates
Type of sampling	Pump sample, scoop sample, tap sample, mixed sample, qualified random sample
Sample vessel	Glass, closure, pre-treatment of the containers, Which materials the sample has come into contact with
Weather	Sun, precipitation, or data from weather services
Blank samples	Field blank sample, transport blank sample
Meta information (depending on the analytical question)	Special features during sampling, e.g., use of a power generator, fertilization in the immediate vicinity at the time of sampling An environment of the sampling point, e.g., abnormalities such as discharges, production facilities, agricultural activities, flooding

7 Reagents

7.1 General information

Special purity requirements must be placed on all reagents used. The content of impurities that contribute to the blank value must be as low as possible or negligible with the signals relevant to answering the question. This must be checked regularly (see chapter 5).

7.2 Eluents

Solvents (e.g., methanol, acetonitrile) and water must be suitable for HPLC and mass spectrometry. Special qualities are commercially available. If the necessary glassware cannot be baked out (see chapter 8.1), it should be easy to clean and exclusively for screening to avoid cross-contamination.

7.3 Operating gases for mass spectrometers

The operating gases for the mass spectrometer must meet the minimum requirements of the device manufacturer. This also includes the materials of the gas lines and gas cleaning cartridges upstream of the inlet of the corresponding mass spectrometer.

7.4 Reference substances

Reference substances are required to safely identify compounds (see chapter 10.2.1). They should have a purity content of at least 95 %, if possible. Solutions of several reference substances (multicomponent standard) can also be used to monitor the stability of the LC-HRMS system (see Appendix E).

7.5 Internal standard substances (IS)

Isotope-labeled compounds should be used (see appendix B.1). They are used to check the measurement stability of each sample and provide possible indications of matrix effects. For example, the IS can be added automatically to each sample via co-injection with the autosampler (e.g., 95 µL sample + 5 µL IS).

7.6 Production of the solutions

When preparing solutions, each work step must be checked for possible contamination. If possible, contact with plastic materials should be avoided. The use of (pyrolyzed) glass syringes has proven to be advantageous in practice.

7.6.1 Stock solution (reference substances)

Stock solutions should be stored at a maximum of -18 °C, protected from light and evaporation. Under these conditions, a shelf life of at least 6 months can generally be expected.

7.6.2 Doping solutions (IS)

Preparing the doping solution in a concentration adapted to the respective substance's detection sensitivity makes sense. This ensures the detection of the internal standards while avoiding overdosing. Overdosing can lead to signal suppression of components contained in the sample during ionization.

7.6.3 QS standard (control standard)

For this purpose, a multi-component standard should be used, and the substances contained should cover the LC-HRMS method's mass spectrometric and chromatographic range as comprehensively as possible. A multi-component standard should also be analyzed in a real matrix to test the generic (generally valid) peak finding. In the best case, this reference matrix should be an aliquot of a representative environmental sample available in sufficient quantity (possibly spiked). This also expands the substance spectrum to include unknown substances at various concentration levels. In addition to checking the recovery and signal intensity of the spiked targets, the intensity-dependent peak-finding parameters used (see chapter 10.1.2) can also be monitored to optimize false-positive and false-negative results.

8 Devices

8.1 General information

Devices or device parts that come into contact with the water sample should be free from residues that could cause blank values. Only glass appliances should be used if possible, as they are particularly easy to clean by baking them out (see also chapter 5).

8.2 Sample vials

Depending on the sample introduction system, rolled rim bottles with a nominal volume of 1.5 mL and crimp caps with septum can be used. Heating the bottles for at least 4 h at 450 °C has proven effective. The cleaned sample vials must be protected from contamination until they are used. This also applies to the sampling tubes. As it is impossible to bake out crimp caps and septa, septa made of materials with the lowest possible blank value entries should be used. For example, PTFE-coated septa should be used in preference to rubber septa.

8.3 High-performance liquid chromatography

8.3.1 General

HPLC systems that are to be used coupled with mass spectrometers for screening purposes usually consist of a degassing device, a low-pulsation analytical pump system (suitable for binary gradient elution), a sample introduction system (optimally cooled for gentle sample storage until measurement) and a device for thermostating the separation column.

8.3.2 HPLC column

Based on the task, the expected variety of relevant substances, and the blank value requirements for detection (data quality) with high-resolution mass spectrometry, HPLC columns must

be selected that have sufficient retention with the use of MS-compatible eluents (organic solvents and possibly volatile buffers).

In addition to reversed phases (RP), typically C18 or polar modified C18 materials, columns with other separation mechanisms (such as HILIC or mixed mode materials) can also be used. HILIC is an MS-coupled alternative to normal phase chromatography for separating strongly polar compounds. It consists of a polar stationary phase (similar to normal phase chromatography, partly in combination with cation/anion exchanger functions) using conventional RP eluents (water, methanol, acetonitrile). The requirements, such as purity and suitability of eluents and ionization aids, must be met for the high-resolution mass spectrometer to ensure high data quality. Examples of measurement methods used are listed in appendix C.1.

To check the robustness, reference materials (e.g., the substances in the IS) that cover the entire separation range should be measured regularly. Similarly, reference substances can be used for standardization, i.e., creating the so-called retention time index RTI (Table B.3), which allows for comparing the retention times between laboratories [24].

8.4 Mass spectrometer

8.4.1 General

The HRMS mass analyzers most commonly used in routine applications today are time-of-flight mass spectrometers ((Q)TOF) and Orbitrap systems (see Appendix D). In isolated cases, Fourier transform ion cyclotron resonance mass spectrometers (FTICR-MS) are also used in water analysis. The mass spectrometers are typically operated for NTS in a combination of full-scan measurement with subsequent measurements in MS^2mode with automated recording of fragment ion spectra (DDA, DIA). Since cycle times are in the ms range, one full-scan MS spectrum and several MS^2 spectra can be recorded per cycle. When preparing the MS method, it is essential to ensure that the cycle times are chosen sensibly concerning the chromatographic peak widths to ensure a sufficient sampling rate. Usually, the measurements are carried out in a specific (positive or negative mode), so two measurements are necessary to completely detect all ion species. Newer HRMS devices also allow simultaneous acquisition (so-called polarity switching) without excessive time losses due to switching times between the two measurement modes. Schematics and explanations of QTOF and Orbitrap systems are listed in Appendix D. Examples of MS methods are listed in appendix C.2.

Specific minimum requirements for basic device parameters are necessary to be able to carry out screening measurements using LC-HRMS:

- The ***resolution*** should be at least 10,000 [25, 26] (at 10 % valley height). This is roughly comparable to 20,000 when calculating the resolution using the half-width of the mass peak.
- The ***mass range*** should be selected depending on the questions asked of the samples. In environmental analysis, most molecules of interest are in the range of *m/z* 100 to *m/z* 1200.
- The ***mass accuracy*** at *m/z* 200 should be at least 5 ppm [26, 27] to limit the number of possible sum formulas. Also, due to older HRMS systems, a practical value in routine is < 10 ppm. Further technical improvements can also achieve a mass accuracy of < 5 ppm in routine applications. The mass accuracy should be checked by regular calibration.

- Various recording modes for ***fragment ion spectra (MS²)*** described inTable 9.3 are possible. The basic definitions of HRMS should also be fulfilled in the MS² spectra.
- The required ***sensitivity*** depends on the task and the chromatography used and should allow the detection of substances in the range of approx. 10 pg absolute. An injection volume of 100 µL corresponds to a content of 0.1 µg/L in the sample. When analyzing water samples, detection limits in the lower ng/L range are required.
- ***System sensitivity and mass accuracy stabilit***y must always be ensured (see Figure E-3).

8.4.2 Ion source

The choice of ion source depends on the task at hand. Electrospray ionization (ESI) has proven to be the best choice due to its universal and robust application. Other ionization techniques (e.g., APCI) can be used analogously depending on the problem or the substances to be detected.

9 Implementation

9.1 Sample preparation

Sample preparation depends on the task, the type of water sample (e.g., leachate, wastewater, surface water, groundwater, drinking water), and the sensitivity of the available LC-HRMS system or the necessary detection limits. To avoid blank values due to contamination (see chapter 5), the aim of sample preparation should be to carry out only essential steps and to be aware of every source of contamination [28]. Table 9.1 contains examples of various sample preparations and sample tasks.

Table 9.1: Advantages and disadvantages of individual steps in sample preparation and sample application

Designation	Procedure/Parameters (Example)	Advantages	Disadvantages
Sample preparation			
Filtration	Membrane filter made of regenerated cellulose, cellulose acetate, PTFE, or glass fiber	Homogeneous sample	Contamination, sorption, labor- and material-intensive, clogging
Conservation	Cooling (4 °C, -18 °C), stabilizer	Acts differently on different analytes differently (may need to be checked in long-term studies)	
solid phase extraction (SPE)	Sorbent material and quantity, pH value, solvent	possibly high enrichment factor, separation	Contamination, sorption, substance groups specific, labor- and material-intensive
Centrifugation	min. 2500 x g, 10 min	Simple and fast implementation	Risk of breakage, contamination, and absorption in the event of decanting
Sample injection			
Direct injection, without SPE	usually 10 – 100 µL	unchanged sample, small sample volume required, automation, little space is required for retained samples	poor limit of quantification
Co-injection of internal standard (IS)	e.g., 95 µL sample and 5 µL IS	Time and cost saving, high reproducibility	not possible with every autosampler
Offline/Online SPE	Sorbent material and quantity, pH value, solvent	Complete automation is possible, with a reasonable limit of quantification	Contamination, sorption, substance groups specific, material-intensive

9.2 Chromatography

The chromatographic separation must not be neglected despite the selective HRMS. The retention time (RT) is essential for identifying a substance and reflects physicochemical properties (e.g., polarity). Which chromatography is used depends on the task at hand. If the separation performance of a classic C18 HPLC column is insufficient, polar-modified C18 materials and/or column materials with smaller particle diameters (e.g., UHPLC columns) can be used. The stationary and mobile phases must be adapted to the polarity range (log D) to be detected.

Additional requirements arise for chromatography as a result of the task. MS-compatible, volatile, and evaporable buffers or ionization aids must be used for separation. The reproducibility and stability of the separation are critical here to make meaningful comparisons within and between different data sets. The comparison of chromatograms, for example, a time series over several months, requires high long-term stability (see Appendix E and E.2). An RT tolerance of 0.15 minutes (analogous to [27]) can be defined as the minimum requirement for RT stability. The retention times can be checked using reference materials. On the one hand, this makes it possible to determine the robustness of the separation and, on the other hand, to standardize the covered separation range (in terms of polarity). This retention time standardization via an RT index (RTI) can enable the transferability of results between laboratories with different LC methods in screening (for an example of an RT standard, see Table B.3). These RT standards can be used as reference values for the so-called 'alignment' of results from long-term studies.

9.3 Mass spectrometry

9.3.1 Ion source / ionization technology

For applying NTS in water analysis, using an electrospray ionization source (ESI) has proven to be the preferred ionization technique. NTS requires an ion source that detects a broad polarity range of analytes with sufficient sensitivity. The source parameters (e.g., temperature, gas flows, voltages) must be selected during ionization so that few fragmentation reactions (in-source fragmentation) or adduct formation occur in the source. Despite the relatively gentle ionization in ESI, fragment formation in the source can rarely be avoided entirely. Alternatively, other ionization techniques, such as APCI, may be helpful depending on the problem or sample. Table 9.2 lists typical adducts and fragments that can form during electrospray ionization. For a more detailed list of typical adducts and fragments, including examples of substances, please refer to Appendix G.

Table 9.2: Adduct and fragment formation in the source during electrospray ionization

	ESI+	ESI-
Substance properties	Sufficiently basic compounds that accumulate protons or other cations (in the gas phase)	Sufficiently azide compounds that release a proton (in the gas phase)
Ionization	Addition of cations	Release of a proton or addition of an anion
typical adducts	$[M+H]^+$, $[M+Na]^+$, $[M+K]^+$, $[M+NH_4]^+$, $[M+nH]^{n+}$	$[M-H]^-$, $[M+ HCOO]^-$, $[M+Cl]^-$, Dimers
Fragmentation	Gentle ionization and, therefore, relatively few fragments (in-source *fragmentation* not always easy to detect)	
Typical fragments	$[M+H-H_2O]^+$, $[M+H-CO_2]^+$, $[M+H-C_2H_6O]^+$	$[M-H-CO_2]^-$, $[M-H-HF]^-$

9.3.2 Measurement technology

The aim of Suspect-Target and NTS is to obtain as much analytical information as possible about the sample during the LC-HRMS measurement. Different measurement modes can be used for this, depending on the task. The measurement techniques are summarized in Table 9.3. In addition to recording high-resolution mass spectra, a predefined or automatically triggered MS^2 spectra recording can be carried out depending on the scanning speed of the device used (see Figure *9-1*Figure 9-1*:* Schematic representation of various possible MS2measurement modes

The mass spectrometric data acquisition (one full scan spectrum including MS^2 spectra per cycle) must be selected to guarantee a sufficiently good image of the chromatographic peaks. The total duration of a mass spectrometric measurement cycle (cycle time) must be adapted to the chromatographic method. A sampling of peaks should be at least 12 data points for robust evaluations [27]. A lower sampling rate can be accepted to capture more information in qualitative screening measurements. However, at least six data points are also required here, as otherwise, the increase in measurement fluctuations makes a reproducible evaluation difficult or impossible.

Table 9.3: Compilation of the different MS measurement techniques with a brief description [29]

Measuring mode	Description
Full scan HRMS	Detection of the accurate masses of all ions formed in the ion source in a specified mass range over the entire chromatographic run time.
MS/MS	Selection and fragmentation of an ion (precursor ion) and detection of the accurate masses of the fragments formed. The precursor ion was selected according to various criteria:
MS/MS Target	Specification of specific masses from which an MS/MS is measured.
DDA or [29]	The device scans for precursor ions over the entire measurement time, and as soon as the threshold value of signal intensity is exceeded, MS/MS fragmentation starts (for example, in Appendix D).
[29]	Permanent/alternating fragmentation of all ions; Option for fast scanning of successive mass windows (MS^E, SWATH®) or fragmentation of all ions (MS/MS^{all} or AIF). Manufacturer-specific, not possible with all devices. Significantly more complex data evaluation (example in Appendix D).

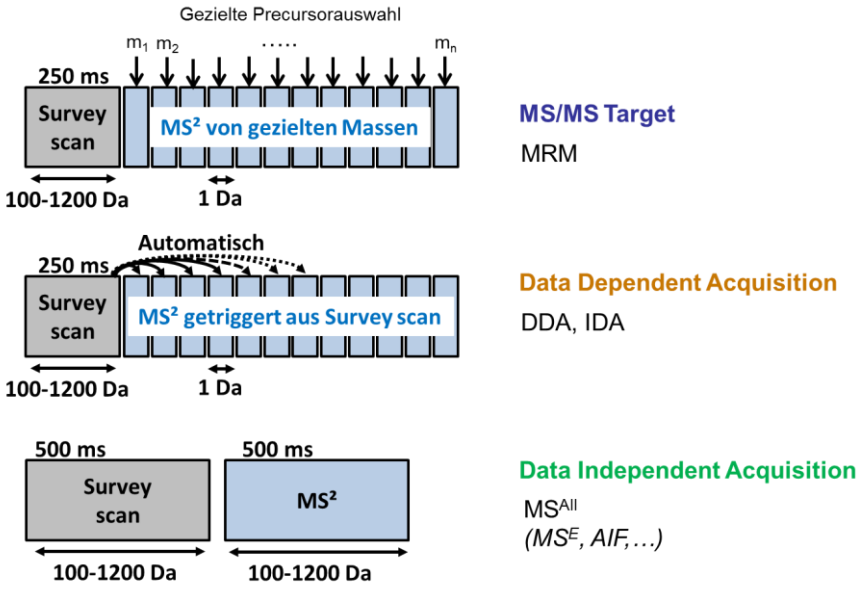


Figure 9-1: *Schematic representation of various possible MS^2measurement modes*

9.3.3 Mass calibration and mass accuracy

The mass calibration must be performed and/or checked regularly and documented depending on the measuring system used. All measuring modes (MS and MS^2) and ionization modes (ESI positive and negative) must be calibrated following the manufacturer's specifications. The prescribed calibration solutions or standards must be used. This mass calibration can be performed internally and/or externally and must cover the relevant mass range.

9.3.4 QA of the LC-HRMS measurement

The use of isotope-labeled substances (see chapter 7.5) as internal standards, preferably distributed over the entire retention time and mass range, is necessary to control system stability concerning retention time, mass accuracy, and sensitivity (possibly taking ionization efficiency

into account) and detection of matrix effects [21]. QA also includes the control of carryover, e.g., through regular system blank values and ensuring sample identity (sample sequence in the autosampler).

10 Evaluation

An LC-HRMS measurement usually results in large measurement data sets of often several million data points. A retention time (*RT*), a mass-to-charge ratio (*m/z*), and a signal intensity (*I*) are stored for each data point. The evaluation of such complex, multi-dimensional data can be divided into two basic steps:

- The data points belonging to a feature (chromatographic and mass spectrometric peaks) are summarized. These features are compiled in so-called feature lists.
- The feature lists obtained are then analyzed or interpreted, e.g., by identifying individual features using a reference database or comparing different feature lists based on a similarity analysis.

In general, various methods are available for processing and analyzing NTS data. Therefore, comprehensive documentation of the respective approach is strongly recommended for traceability and transparency. The evaluation platform used or the user input parameters set, such as intensity filter thresholds, etc., can be recorded to ensure that the results from the measurement data can be reproduced. Changes to the processing and data analysis strategy can make it difficult to compare results and should, therefore, always be checked for accuracy and plausibility, e.g., using reference standards.

10.1 Processing the measurement data

The device manufacturer's software is generally used to evaluate the LC-HRMS raw data. This can be supplemented or replaced by software from other manufacturers or in-house developments, depending on the problem. In addition, numerous algorithms have already been developed, which can also represent advantages over the respective individual solutions. Examples of larger platforms and projects include patRoon [30], MZmine [31], EnviMass [32] and XCMS [33]. A major advantage here is the comparability of the evaluation despite the use of different device technologies. After converting the acquisition data into free formats, for example, *.mzML or *.mzXML, a wide variety of data formats can be processed using the same workflow.

The first steps of data processing are decisive for the result of the NTS [34] and will be explained individually and in more detail below.

10.1.1 Centroidization

High-resolution mass spectrometry can precisely determine the *m/z* of most ions. However, this requires that each *m/z* be recorded in a narrow range of values over many measuring points. A normally distributed peak profile is usually detected for the respective m/z if a corresponding ion is present. This profile data is usually converted to the actual *m/z* lines by centroidization. It is essential to differentiate between the term "peak," which is used both for the high-resolution mass spectrum profile data and in the chromatographic profile data explained later. Therefore, an appropriate classification in the respective context is essential to avoid confusion.

Due to technical fluctuations such as signal intensities, the determined *m/z* lines (centroids) vary slightly within an overall measurement and should always be considered in combination. There are two different methods for determining the centroids, the results of which may differ.

- Local maximum: Here, the centroid is equated to the data point with the maximum intensity in a measured peak profile, i.e., *m/z* and intensity are taken directly from a measurement point.
- Interpolation using a distribution model: A theoretical distribution curve (usually normal distribution) is adjusted based on the measured peak profile, and the center of gravity is determined from this theoretical curve.

To illustrate this, both variants are shown in Figure 10-1. The simplicity and speed of the method speak in favor of using the local maximum, while the accuracy can suffer in the case of overlapping peaks and noise. Centroidization via a peak model offers higher accuracy, especially for more complex spectra, but requires more computing power and a suitable model parameterization. However, there is no guarantee that these methods are error-free. Therefore, comparing results by reference substances is essential to ensure reliable and accurate mass spectra analysis.

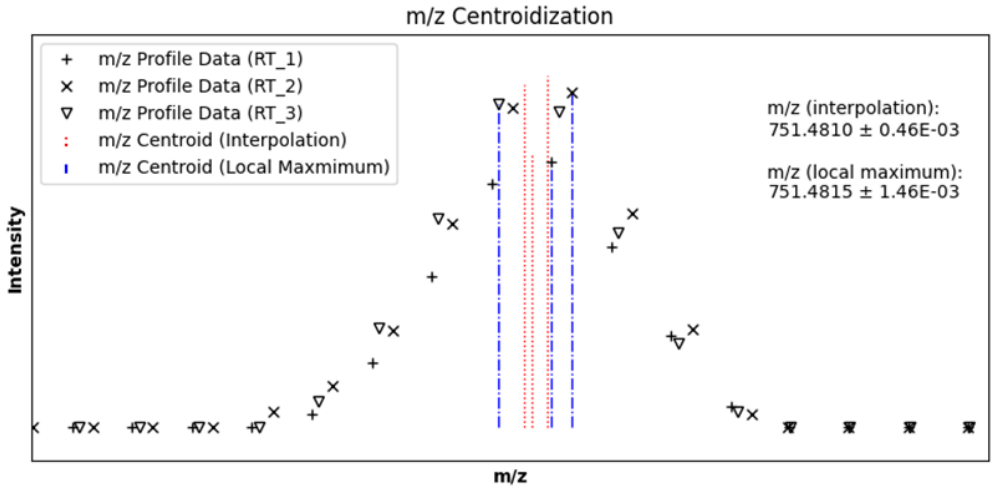


Figure 10-1: *Extract of three consecutive high-resolution mass spectra in the range m/z = 751.48. The three data sets each show a typical measurement signal with bell curve progression. The maximum of this peak profile is estimated when the center of gravity is formed. A comparison of two estimation methods (interpolation and local maximum) reveals subtle differences.*

The first reduction in the file size can already be made by centroidization. Depending on previously defined criteria such as intensity or signal-to-noise ratio thresholds, not every peak profile is converted into a corresponding *m/z* line. With some measuring instruments, the centroidization is formed before the raw data is saved so that, in these cases, only the already compressed or calculated data sets are available for evaluation.

10.1.2 Peak finding

Determining features is an important step in preparing the measurement data, on which all further evaluations depend. Generally, a feature is a chromatographic peak of an *m/z* mass trace that fulfills a previously defined criterion, e.g., intensity threshold value at the peak maximum. Typically, a feature is summarized by the properties *RT*, *m/z,* and *intensity* already described and can be supplemented by additional information such as an MS^2 spectrum. When

determining the *m/z* values of a feature, the scatter of the data points in this domain must be considered. Various methods exist, such as the average value or the weighted mean.

Depending on the task, peak finding or feature detection can be performed manually, for example, using a suspect target list. With NTS, peak finding is carried out using special peak finding algorithms. There are different strategies, three of which are listed below as examples:

- The first strategy considers the two coordinates, RT and *m/z,* independent. The fluctuation of a mass is considered via the *m/z* axis, and the course of intensity is considered via the retention time axis. The definition of an intensity threshold is a decisive criterion for feature detection.
- The second strategy analyzes extracted ion chromatograms within a narrow *m/z* range. These chromatograms can then be analyzed independently using a suitable filter (e.g., second-order Gaussian filter) for chromatographic peaks. This strategy avoids the search for peaks in the complete *m/z* range.
- The third strategy for extracting features from an LC-HRMS data set is to apply modeling to the raw data. This approach uses a three-dimensional model where RT and intensity represent the peak profile (e.g., a Gaussian model), and *m/z* represents the isotopic pattern (e.g., the 13C isotope). The fit starts with the most intense peak of the entire data set and is then subtracted. This process is repeated until only background noise remains.

For further details, please refer to [35].

Regardless of the peak finding strategy used, input parameters are usually required for the corresponding algorithms, such as an intensity or signal-to-noise ratio threshold value or the width of the *m/z* range *to* be considered.

As a result of peak finding, a feature list is generated, which is subsequently analyzed and interpreted. However, it should be noted that accurate and robust feature detection is one of the significant challenges of NTS measurement data processing. In this context, for example, false-positive entries may end up in the feature list, or actual features may not be recognized (false-negative). This problem should be taken into account when optimizing all peak-finding parameters. False negatives are more likely to occur if the criteria are too strict, i.e., real signals are no longer detected automatically. On the other hand, settings that are too generous increase the false-positive rate by detecting noise that is incorrectly recognized as a peak. This opposing behavior of false-positive and false-negative findings complicates optimizing peak findings and requires compromises depending on the question. It may make sense to minimize the number of false-negative findings and initially accept an increased false-positive rate. This can be reduced by additional filter criteria (after the actual peak finding). To determine the false-negative rate, for example, known standard substances can be added to check whether they can be found in the feature list. To determine the false-positive rate, the chromatographic profiles of randomly selected features can be evaluated visually, e.g., concerning the peak shape. However, it should be noted that evaluating peak profiles requires a great deal of experience and always includes subjective impressions. An evaluation by several specialists can be helpful [36]. An example of false-positive chromatographic profiles can be seen in Figure 10-2. Automated filters, e.g., maximum intensity, mass and/or time deviation of features between replicate samples and minimum peak quality parameters such as signal-to-noise ratio, peak time, mass width, etc., can be applied to further reduce the number of false-positive features [37]. Most peak-finding algorithms already apply some of the mentioned filters but often still generate low-quality features to avoid false negatives. Therefore, applying these filters to

the extracted chromatographic peak of a particular feature and to the correspondence of features between replicate samples can prevent the inclusion of these false-negative features.

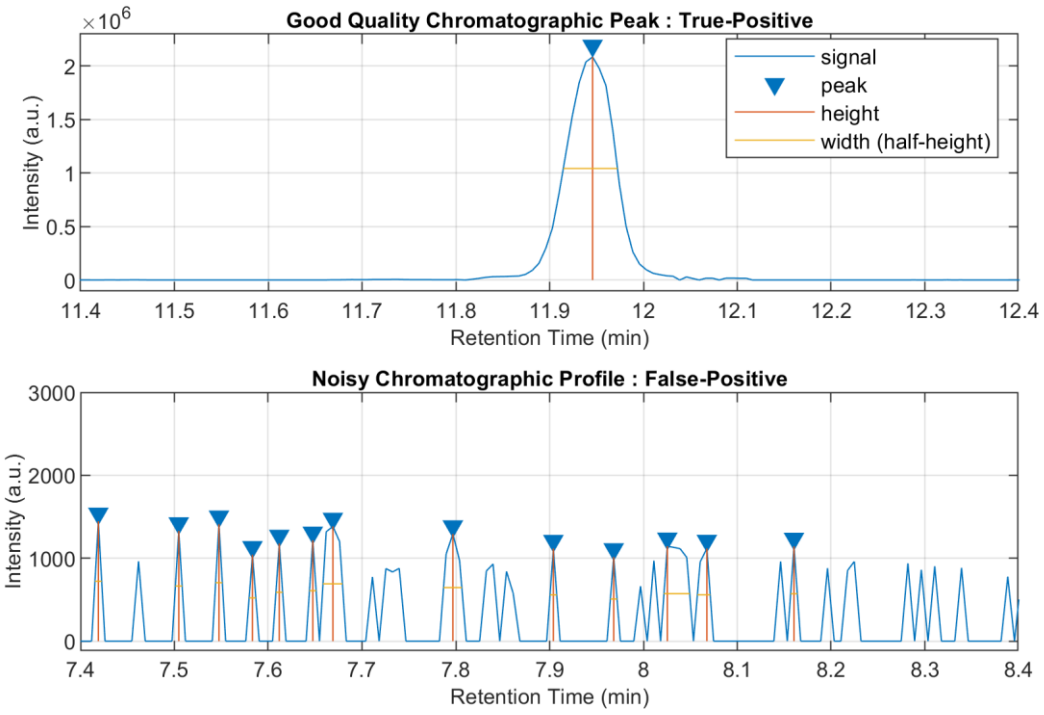


Figure 10-2: *Examples of correctly (top) and incorrectly (bottom) detected peaks in chromatographic profiles. Many of the incorrectly detected peaks shown below can be filtered out by suitable settings of the respective peak detection algorithm.*

The so-called intensity threshold greatly influences the result, which defines the signal level up to which features are to be searched for or which data points are to be taken into account for determining the features. Since a feature is determined from a series of data points and is not based on individual measurement points, the intensity threshold can influence the RT, *m/z,* and intensity properties, primarily if these are based on the peak area. The intensity threshold should be optimized so that most known components in the relevant concentration range can still be detected. The verification or validation of the peak finding is described in Chapter 11.1.

10.1.3 Alignment

Detected features can fluctuate slightly in their properties, such as RT and *m/z*, so the features of the same substances can differ slightly over several measurements. In feature alignment, these fluctuations are considered by previously defined tolerances to ensure the comparability of several samples (e.g., a time series). Different software packages use different terms for this sample merging step.

Alignment involves assigning the same features within and between different samples. The alignment of the detected features takes place in the retention time and mass domains. The result can, for example, be a data matrix consisting of features (rows) and samples (columns), with the peak intensity or peak area as a matrix entry. To improve the assignment between the samples, a retention time correction and mass recalibration can be carried out, for example, using internal standards (see appendix B.1)

10.1.4 Blank value correction

When processing data, it is crucial to consider the blank value. First and foremost, this should minimize false-positive findings. To this end, the blank value must be selected to match the samples. If the incorrect blank value is included in the evaluation, there is a risk of eliminating real features (generation of false-negative findings). For example, a blank value in a system, field, or transport is helpful for the direct measurement of samples. When measuring prepared samples, such as SPE extracts, false-positive findings are minimized by selecting an extraction blank value. More detailed explanations of possible blank values and their consideration are given in chapter 5. There are various approaches for handling the blank value correction. All features found in a blank value sample can be deleted, or their intensities/peak areas can be subtracted from the intensities/peak areas in the samples.

An example is the matrix correction approach [37], which considers the suppression by the matrix in the sample compared to the corresponding blank. The correction of the matrix suppression improves the comparison between the samples. In most evaluation programs, selecting only signals that exceed a certain intensity ratio compared to the blank value is also possible. Typical values here are e.g., 3-fold or 10-fold higher intensity than the blank value.

10.1.5 Componentization

A molecule can generate different adducts during ionization (see Appendix G). The isotope pattern is also available for each of these adducts. In addition, fragmentations can occur in the ion source that generates further features for the molecule (see chapter 9.3.1). In this way, numerous features can be assigned to a molecule under certain circumstances. Through componentization, these features can be identified and combined into one component. The terms used for these combined components vary depending on the software package and device manufacturer (e.g., Molecular Feature (Agilent), Bucket (Bruker), Feature (Sciex), Compound (Thermo)).

Algorithms based on predefined rules and logic are typically used for componentization. These algorithms are often an integral part of the device software and support identifying and grouping features. Alternatively, open-access software can also be used here. For example, the CAMERA package [38] for the R programming language uses chemical information such as isotope distributions and fragmentations to define components. Another method is Multivariate Curve Resolution (MCR), which is based on statistical criteria and determines standard elution profiles of different masses. Such approaches enable a differentiated and precise analysis of the data, which is particularly important in NTS (for further details, see chapter 10.2.1.2 and Appendix H

10.1.6 Sum formula generation

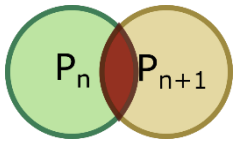
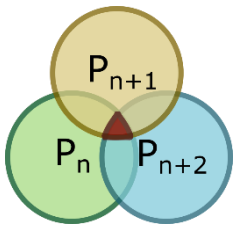
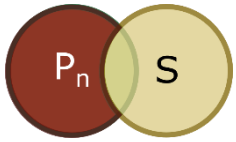
Possible sum formulas can be specified based on the determined accurate mass and the isotope pattern. Additional information from any available MS^2 spectra can also increase the certainty of the sum formula results. The *"Seven Golden Rules"* for determining sum formulas from measurement data are described in [11]. The more precise the mass is determined, the fewer possible summation formulas are. The type and scope of the summation formula suggestions also depend on the elements used to calculate the summation formula. A clear summation formula can rarely be obtained from the measurement data [11]. However, it should be noted that the *"Seven Golden Rules"* method may have weaknesses when analyzing unusual and complex compounds, such as PFAS, as such compounds may have atypical isotope patterns, unusual bonds, valences, and deviating elemental ratios.

When identifying unknown features, it should be noted that some substances do not form simple $[M+H]^+$ or $[M-H]^-$adducts; these only make up a small proportion of the detected ions. Furthermore, some substances undergo rearrangements and/or decomposition reactions during ionization, making identification even more difficult.

10.2 Interpretation

A prerequisite for answering the question (Table 1.1) is validated data from the evaluation (see chapter 10.1). The results can be displayed in a mass-retention time plot, for example. The entirety of the features can be viewed as a set P_n (in the mathematical sense). The features (components) are characterized by the accurate mass and the retention time. The quantities can be evaluated comparatively according to the question. For example, some questions relating to a temporal sample series are summarized inTable 10.1 with the symbolic representation and the formulation in the quantity notation.

Table 10.1: Schematic representation of the comparison of features between samples

Question	Symbolic representation	Set-theoretical description
The feature is included in two consecutive samples		$\boldsymbol{P_n \cap P_{n+1}}$
Feature is included in three consecutive samples		$\boldsymbol{P_n \cap P_{n+1} \cap P_{n+2}}$
Feature is included in all 14 samples in the series	$\boldsymbol{S = P_1 \cap P_2 \cap \ldots P_{14} = \bigcap_{1}^{14} P_i}$	
Feature is only included in one sample of the series		$\boldsymbol{P_n \setminus S}$

10.2.1 Identification

Depending on the information available, it makes sense to classify the degree of identification into categories or levels [39]. Uniform categorization is a prerequisite for comparing results from different laboratories. It is helpful to distinguish between two groups of addressees when communicating the results from the NTS. One group represents addressees without detailed knowledge of measurement technology and evaluation, while the other group has this detailed knowledge. The purpose of this differentiation in communicating results is to focus on the information essential for the addressee. Table 10.2 shows the classification with the corresponding requirements.

The categorization is based on the information generated using LC-HRMS: Retention time, accurate mass, and the measured MS² spectra. Additional measurement data, such as the ionization mode and CCS (Collision Cross Section) values from ion mobility measurements, can further contribute to narrowing down database hits and unambiguous substance identification [40].

Table 10.2: Classification of features from the HRMS screening (based on Schymanski et al. 2014 [39])

Client Protocol		**Protocol Editor**						
				Reference Data				
Class	**Statement**	**Signal***	**Statement**	**accurate mass**	**RT (RTI)**	**MS2 database**	**MS2 reference**	**MS2 *in-silico***
Cat. 1	Identified substance	**Cat. 1**	Confirmed substance/structure	✓	✓	✓	✓	!
Cat. 2	Probable substance	**Cat. 2*****	Probable substance/structure	✓	!	✓	✗	!
Cat. 3	Substance proposals from sum formula	**Cat. 3a**	Possible structure, specification of metadata	✓	!	✗	✗	✓
		Cat. 3b	Possible substance	✓	✓	✗	✗	✗
Cat. 4	Signal of a substance	**Cat. 4a****	Sum formula	✓	✗	✗	✗	✗
		Cat. 4b	Feature (signal)	✗	✗	✗	✗	✗

*A signal is characterized by the exact mass, the retention time, and the intensity.
**A sum formula can be given if at least two isotopes and/or adducts can be assigned in the signals. Depending on the problem, one isotope may also be sufficient.
***Confirmation by reference is pending.

Legend:

✓ must be present
! may be present
✗ not available

10.2.1.1 Databases

Using databases to support the identification of features is a fast and efficient procedure. MS spectra databases and substance databases may be available for a fee. Examples of MS spectra databases are MassBank [12, 41, 42] and mzCloud [13, 43]. Success depends on the search criteria and the scope of the database entries. A large number of different substance databases are available on the Internet. There may be hundreds to thousands of hits for a searched mass or molecular formula for general chemical databases with several million entries, such as PubChem [14] and ChemSpider [15]. The databases sometimes allow multiple hits to be prioritized using meta-information. For example, a retention time estimate using quantitative structure retention models can help prioritize structure proposals matching the measured retention behavior [44]. Other meta-information that can be used to prioritize hits includes literature references, toxicity data, or intended use and quantity. The FOR-IDENT work platform [18] with the stored substance database STOFF-IDENT [17] and other environmentally relevant substance databases such as CompTox Chemistry Dashboard [19] and Norman Network Databases [45] provide support in the identification of substances relevant to water in particular, where the database is not only queried for accurate mass or molecular formula, but additional information (metadata see 10.2.1.2) can also be included. For example, the FOR-IDENT platform uses the normalized retention time, sum formula and/or production spectra (comparison with in-silico fragment ion spectra) to prioritize an individual proposal in the event of multiple hits for a queried mass or sum formula.

10.2.1.2 Metadata

Information about the sample under investigation helps identify features or components and match these with substance properties. Such metadata include, for example, occurrence, areas of use, quantities used, possible transformation or by-products from production or their application.

10.2.2 Statistical methods

The findings are often placed in a higher-level context as part of evaluating NTS data. The comparison of different samples or with corresponding references usually plays an important role here. Various approaches to data analysis are available here, including univariate or multivariate statistical methods (see Appendix H) [34, 46]. These methods can either be supervised approaches, where known annotations are used as a reference, or unsupervised, in which case no prior information about the annotations is available. Both approaches examine the data structure and select relevant variables (features).

Example methods: When comparing two groups (e.g., effect/no effect or different process stages of a wastewater treatment plant), relative intensity changes, also known as "fold changes" (f_c), can be determined. These are usually specified as a logarithm to the base two ($\log_2$) to enable a symmetrical representation of up and downregulation. Thus, an increase or decrease in the $\log_2$fold change by 1 means a doubling or halving of the intensity of the respective feature.

$$\log_2 f_c = \log_2 \frac{feature\ intensity\ sample\ A}{feature\ intensity\ sample\ B}$$

The Volcano plot is a particularly informative tool for visualizing and prioritizing individual features using fold changes. In a Volcano plot, the $log_2 f_c$ is plotted on the x-axis against the

negative logarithm of the p-value ($-log_{10}(p)$) on the y-axis. The p-value indicates the probability to what extent the examined differences between the two groups are due to chance alone. This representation makes it possible to quickly identify and highlight features with significant intensity changes. Features that lie far to the right or left of the center (high $log_2 f_c$) and at the same time far above (low $-log_{10}(p)$-value) in the plot are particularly conspicuous, as they show both a significant change in intensity and a high statistical significance. This makes the Volcano plot an efficient tool for prioritizing features within the NTS, especially when highlighting potentially relevant connections in complex datasets.

Unsupervised methods such as principal component analysis (PCA) can be used to explore data or identify relationships between samples. This technique makes it possible to identify the essential variables (features) that have a large share of the total variance across all samples, as seen in the loading plot by large amplitudes. In addition, correlating features can be identified and grouped using PCA as they have similar loading profiles.

Another vital aspect of PCA is the evaluation of the reproducibility of measurements, particularly by considering multiple measurements of a sample. In the PCA scores plot, multiple measurements should have similar characteristics and be close together in the graphical representation. This is an indicator of the consistency and reliability of the measurements. If multiple measurements are not close together, this may indicate possible outliers, measurement errors, or other extremes that require further investigation. The spatial proximity of multiple measurements in the scores plot can be used to assess the quality and reproducibility of the analytical data in the NTS.

Many cluster analyses, e.g., hierarchical clustering analysis (HCA), also fall into unsupervised methods and can be applied to feature lists. The similarity of samples and a grouping is determined based on similar intensity profiles along the feature list and visualized in a dendrogram. Various methods can be used to calculate the similarity of the samples (distance metrics) and to form groups (fusion algorithms).

In general, no groups are defined in advance for unsupervised methods. Groupings can only emerge if these are dominant features in the overall data set under consideration. Supervised methods can be used to check whether and which features are characteristic of a previously defined grouping. Partial least squares discriminant analysis (PLS-DA) is an established classification method. Like PCA, PLS-DA converts the data set into scores and loadings. The difference, however, is that the target value is not the maximization of the variance of all samples but the maximization of the variance of previously defined sample groups. PLS-DA can then be used to determine "Variable importance in projection" (VIP) based on which relevant variables can be selected. In addition to conventional PLS-DA, there are many extended PLS-DA approaches, such as OPLS-DA (Orthogonal partial least squares discriminant analysis) or MOCA (Multi-block Orthogonal Component Analysis), which take into account a wide range of additional criteria to improve the interpretability of the results.

ANOVA-simultaneous component analysis (ASCA) can be used to analyze the influence of several factors of an experimental design on the data set in a model and to estimate the influence of each factor alone and their interaction [47].

Parallel Factor Analysis 2 (PARAFAC2) is specifically designed for analyzing three-way data where one of the dimensions may vary. This makes it ideal for complex data sets where the number of features may vary for different samples. By decomposing the data into factors, each corresponding to one of the three dimensions, PARAFAC2 enables the identification and iso-

lation of consistent patterns and relationships across different datasets. This method is particularly used to uncover hidden structures and facilitate the interpretation of complex data. Its application ranges from identifying unknown contaminants in water samples to investigating changes in water quality over time.

MCR-ALS (Multivariate Curve Resolution - Alternating Least Squares) views the raw data matrix as m/z-based elution profiles, which are broken down into individual components using bilinear regression. Each component represents a calculated mass spectrum linked to a specific elution profile. MCR-ALS does not require prior feature detection and can be considered an alternative evaluation method.

General challenges: The intensities of the different features and their fluctuations across different samples can vary greatly (heteroscedasticity), leading to a disproportionate emphasis on high signals with significant variance. Before a chemometric comparative evaluation, the NTS data should be scaled, transformed, or normalized. A further challenge is that not all analytes or features are present in all samples to be compared. If several samples are merged by alignment, this leads to zeros or gaps in the data matrix. Depending on the software, these can be filled with small intensity values or noise (gap filling, noise filling) or remain as empty spaces.

Another essential aspect of NTS data analysis concerns the required number of samples. It is crucial to analyze a sufficient number of samples to achieve statistically significant and representative results. The exact number depends on various factors, including the samples' variability, the system's complexity under investigation, and the study's specific objectives. An insufficient number of samples can lead to insufficient statistical power and thus affect the reliability of the results. This is particularly true for methods based on comparing groups or aiming to identify patterns in large data sets. In practice, careful consideration and preliminary research are needed to determine the optimal number of samples required to achieve valid and meaningful results within the NTS.

Note on the use of machine learning and artificial intelligence in non-target screening: Machine learning (ML) and artificial intelligence (AI) are promising technologies that can be used in many areas of data analysis in NTS. Although the development of suitable methods in this area is still in its infancy, they offer great potential for improving data collection, processing, and interpretation. A significant obstacle to applying ML and AI in NTS is the acquisition of suitable training datasets, which must contain sufficient high-quality reference data to be effective. A particular strength of ML and AI methods lies in pattern recognition. This makes them particularly useful for applications such as early warning systems and detecting anomalies in complex data sets. Such systems can, for example, help identify unknown contaminants or detect changes in water quality early. From a mathematical point of view, the vast number of features (thousands of features) is usually only matched by relatively few samples, leading to overdetermination and making using AI methods more difficult.

Given the rapid developments in this field, it is advisable to follow the progress and new methods within the NTS community regularly. Integrating ML and AI into NTS can significantly improve the efficiency and accuracy of analyses and provide new insights into complex environmental chemistry issues.

11 Evaluation and reporting of the results - Analysis report

Transparency and minimization of subjectivity in data evaluation are required to generate reproducible and comparable results. Decisions must be made for each step of the data evaluation (setting threshold values, determining false-positive rates, identification, etc.). Clear decision criteria should be defined to reduce the subjectivity of these decisions, and, if possible, evaluation steps should be automated. In addition, the workflow's selected parameters, criteria, etc., should be documented transparently.

11.1 Evaluation of peak finding

As the results of peak finding depend mainly on the parameters to be set, optimization for the respective NTS data sets is essential and should be checked regularly.

To check the peak finding step, doping of known (isotope-labeled) substances in the relevant concentration range (e.g., 0.1 µg/L) in real matrices is recommended for each new measurement campaign (QA sample; see chapter 7.6.3). The intensity dependence of this step requires good long-term stability of the sensitivity of the measuring device (see appendix E.2), which can also be checked. In generic peak extraction, the intensity-dependent parameters (e.g., noise threshold) are decisive and significantly influence the number of features found.

For technical reasons (e.g., adjustment of the detector voltage, replacement of the detector or the ESI needle), the basic sensitivity of the measuring device can vary significantly between two measurement series. Therefore, the values determined for the intensity-dependent parameters during method development and validation must be checked and adjusted regularly. Even when transferring existing evaluation methods to a new measuring system, the sensitivities of the measuring devices will likely differ. The methodology must be adapted; Appendix F provides an example of such a strategy.

The "performance" of the evaluation method can be assessed by validation based on the QA sample. Standard measures such as the false positive rate (proportion of cases where background noise or artifacts are incorrectly identified as positive features), recall (proportion of correctly identified positive features among all actual positive features), or precision (proportion of correctly identified positive features among all features identified as positive) allow a comprehensive evaluation of this step. Recall and precision can be calculated using the following formulas.

$$Recall = \frac{TP}{TP + FN} = \frac{correctly\ detected\ features}{total\ number\ of\ true\ features}$$

$$Precision = \frac{TP}{TP + FP} = \frac{correctly\ detected\ features}{total\ number\ of\ detected\ features}$$

TP stands for True Positives and describes the number of detected features that actually originate from the analyte signals of the QA control sample. FN stands for False Negatives and describes all falsely undetected features, for example, in the control sample. Accordingly, the sum of TP and FN is the total number of features in the QA sample. The sum of TP and false positives (FP), on the other hand, represents the total number of detected features. Figure 11-1 serves as a better illustration in this context.

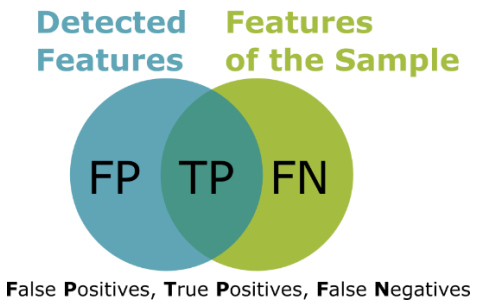


Figure 11-1: *Schematic representation of the different result classifications in feature detection in non-target screening. From an unknown set of actual features in a sample, recognized features can be true (True Positive) or false (False Positive). Unrecognized features are also classified as false (false negatives).*

The formulas described above show that false-positive detected features lead to lower precision values, and false-negative detected features lower the recall value [48]. The main difficulty in determining these performance parameters is determining FN, i.e., not all features in the QA sample's data set have been detected. This assumes that all signals of the QA sample are known precisely. As an approximation, it is therefore assumed that each substance in the QA sample has at least one primary and plausible signal in the NTS data set. The more precisely the number of actual substance signals of the QA sample is known, the more precisely the FN rate can be estimated [49]. The quality of all subsequent steps and the final results are significantly influenced by peak findings, which underline its importance.

11.2 Evaluation of the alignment

From a statistical perspective, the alignment checks whether features with similar properties, such as *m/z* and/or RT, belong to the same population of features and can, therefore, be considered comparatively. Various algorithms are also used for this step, and limit values for mass and retention time windows are defined, which can lead to false-positive or false-negative assignments and should, therefore, be validated. Internal reference standards or a QA sample are recommended for this purpose. False-positive alignment occurs when features of two or more similar substances are incorrectly combined into one feature. False-negative alignment, on the other hand, occurs when features of a substance are not combined.

11.3 Evaluation of the statistical data analysis

NTS datasets usually contain thousands of features (variables) even after comprehensive data processing, so their number usually exceeds the number of samples by far. Due to this problem, also known as "undersampling," the chemometric results should be statistically validated, e.g., by cross-validation or permutation tests [50].

11.4 Evaluation of the database search

The integration of databases, in particular, offers the opportunity to carry out a comparative analysis of the database hits by making the right choice and documenting the parameterization of the query. Furthermore, the uniform description of the analytical certainty (categorization) of identification results of unknown features forms a basis for comparing LC-HRMS screening results (see 10.2.1).

12 Validation

Validation is the confirmation by examination and provision of evidence that the particular requirements for a specific intended use are met [51]. The following key points arise from this description:

- Intended use (specifically intended use)
- (Special) Requirements
- Proof (objective)

These points are also associated with the risk that the intended use of the NTS method does not meet the requirements and is unreliable. This risk must be taken into account when providing objective evidence.

When validating an NTS method, components of target validation must be adapted for NTS. They must be supplemented by measurement parameters and data evaluation (e.g., peak detection, peak alignment) and the software used for data interpretation (e.g., multivariate statistical methods, see Appendix H). The challenge of NTS analysis is considering not just a single sample but several correlated samples. Different statistical evaluation methods must be integrated into the analysis method and included in the validation.

Sampling, sample transportation, and storage of the samples until analysis are part of the analysis procedure and must be considered during validation. A central point is investigating and including suitable blank values during sampling, sample transportation, and sample storage.

It may be necessary to adapt a validation to the task to demonstrate the method's performance for this problem. This means an NTS method can only be validated in general terms (basic validation). Additional validation steps may be required by adapting the method to the problem, e.g., sampling, chromatography, mass spectrometry, and evaluation. Here, validation and quality assurance merge seamlessly [52, 53].

The following points, for example, must also be taken into account during validation:

- Availability of a sufficient number of samples (statistical evaluation)
- Statistical certainty of the result
- Application of suitable experimental designs (DoE) concerning effort and statistical statements
- Dependence of the result (significance) on the laboratory equipment

A prerequisite for validation is a standard operating procedure (SOP). This SOP must be written so different employees can understand and implement it. In the case of an NTS method, this applies to measurement, data preparation, and interpretation.

There are two basic approaches to validation: the single-laboratory approach and the interlaboratory approach (interlaboratory test). Cross-laboratory validation of an NTS method is desirable if it is intended for broad application, such as nationwide water monitoring. The implementation of interlaboratory comparisons (as of 04/2025) is still under development due to the variability in NTS measurement technology and evaluation, for example. The supervisory authority can accept methods (validated in an individual laboratory) for regulatory purposes, depending on guidelines or legislation.

12.1 Validation parameters - Validation plan

Figure 12-1 illustrates the many variables that influence the analysis result. These must be taken into account when validating an NTS method. The influence of sampling and sample transportation on the NTS analysis is a central point that can only be integrated into a validation with great effort due to the variability. Therefore, the validation must include appropriate blank samples for characterization and consideration of sample changes due to sampling and sample transport.

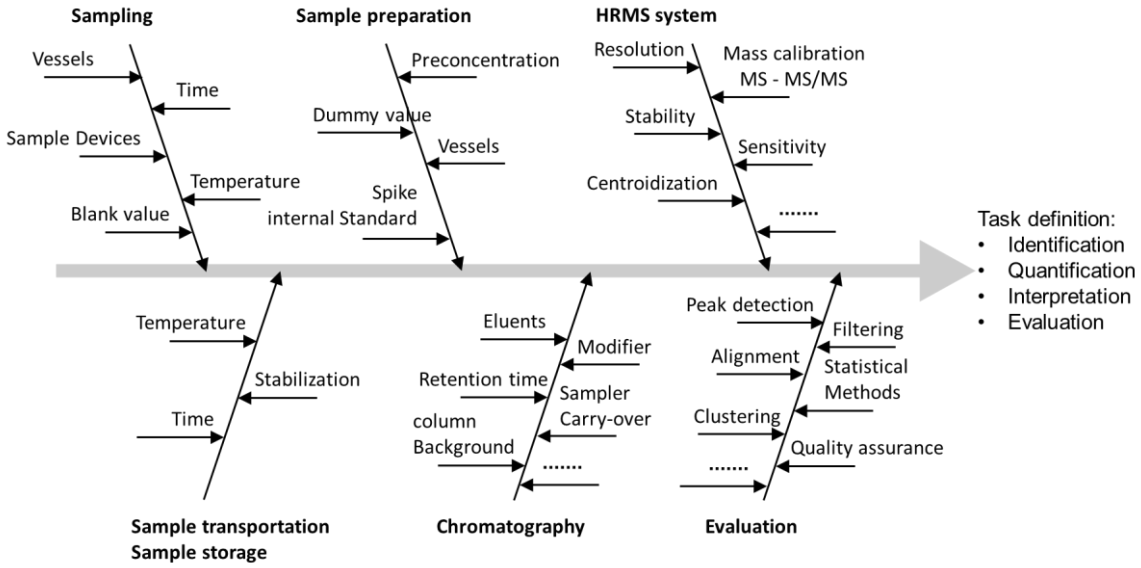


Figure 12-1: Selection of influencing variables on the analysis target (Cause and Effect Diagram)

A validation plan must be drawn up at the start of the validation. This defines the performance characteristics with the performance criteria as far as possible. Performance characteristics and their description are compiled inTable 12.1. Depending on the task, the exemplary performance characteristics that are relevant for a task must be selected from Table 12.1 before the start of validation and documented.

12.2 Qualitative NTS analysis results (classification of samples)

One task of the NTS can be assigning the sample to a specific group (class), such as a treated or untreated wastewater sample. An assignment is made via a sample comparison (peak pattern, feature list) of the samples and not exclusively based on individual identified features. The validation of a classification requires a corresponding number of reference samples. A confusion matrix [53] can be used as the basis for validation. This is not limited to two classes (binary) but can be extended to several classes. Only two- or multi-level results of the NTS method (e.g., treated / non-treated sample, qualitative result) can be used. A confusion matrix can also validate quantitative results divided into two groups (dichotomized) via a selected threshold value (decision value, cut-off). The structure and evaluation of a confusion matrix is shown in Appendix J with two application examples.

Table 12.1: Compilation of performance characteristics for validation

Feature	Description	Performance criteria Example	Implementation / Evaluation
Analytical requirements	Description of the task and the scope of the NTS method	s. Table 1.1	
	Definition of the boundary conditions - Complete validation - Validation due to change in the scope of the application - Validation after changing the process parameters (evaluation method)	Compilation of the performance requirements and justification for their definition	
Selectivity Specificity	Selectivity is the ability of a method to detect different components of a mixture that are to be determined side by side without mutual interference. Specificity is the ability of a method to measure the analyte in a complex mixture in the presence of other mixture components (e.g., other analytes, matrix components) without interference. In the NTS, the relevant (defined) substances (substance groups) in the samples (matrices) must be detected despite possible interfering substances.	Proof that the matrix (interfering components) does not prevent the detection of the substances. The entire workflow (including peak finding) must be considered when checking the performance feature. The validation includes the evaluation algorithms.	Examination of matrix samples spiked with different concentrations. Determination of recovery by comparison with standard solution. The selected relevant substances should cover the applied mass and retention time range. The prerequisite is a reliable identification of the features (see point Correctness).
Detection limit (NG) Limit of Quantification (NTS usually focuses on the reliable detection of compounds).	The lowest concentration of the analyte that can still be detected. The lowest concentration of the analyte that can be quantified with acceptable measurement uncertainty.	Requirement depends on the task, e.g.: - Toxicity of the substances - Environmental concentration - Process behavior - Transformation of substances Evaluation criterion: Signal-to-noise ratio S/N	Examination of matrix samples spiked with different concentrations, taking into account the task. Multiple measurements (at least 3 repetitions) on the NG and determination of S/N (S/N > 3 for all repetitions)

Feature	Description	Performance criteria Example	Implementation / Evaluation
Without suitable reference substances, only a semi-quantification is possible [54]		The entire workflow (including peak finding) should be considered when testing the performance characteristic, as the validation includes the evaluation algorithms.	If no noise can be measured, the NG can be determined using the standard deviation of the measurement signal of a sample with a concentration at the NG. The prerequisite is reliable identification of the features. A fragment ion in MS/MS (independent of DDA or DIA) should be available (see feature correctness).
Work area	The concentration range is a significant dependency between concentration and measurement signal.	The parameters defined in the method (e.g., RT, m/Δm, mass accuracy) should be guaranteed over the working range.	Analysis of matrix samples spiked with different concentrations in the desired working range. The evaluation can be carried out by looking at the point-to-point gradient. In addition to TIC intensities, ion transmission and filling times should also be considered to return to the dynamic range via any dilutions.
Accuracy Measurement uncertainty	- Correct identification of features - Correct classification of samples	Confirmation by reference material Accuracy and reproducibility of mass (MS and MS/MS) and retention time Correct sample classification based on the analysis method and the evaluation algorithms used and evaluation. (see chapter 11)	Specifications for determining the accurate masses of molecule and fragment ions (e.g., from EIC peak maximum, MS/MS, possibly separate measurement). Use of databases. The approach is best defined in separate SOP (see appendix J.2) Carrying out the feature comparison of repeat measurements and determining the

Feature	Description	Performance criteria Example	Implementation / Evaluation
	- Comparison of samples based on the detected features (e.g., Venn diagram). - Quantification (difference between measured value and reference value)	Reproducibility of the features (number, identity, signal intensity) with multiple sample measurements. Quantification of identified substances is only possible after calibration with reference standards.	agreement based on RT, mass, and intensity (see Appendix JJ.1) For quantification, specific separate measurements must be carried out (e.g., by spiking the sample with the identified substances, external calibration, and normalization using internal standards). The approach is best defined in a separate SOP.
Precision Repeatability Laboratory precision	The extent of agreement between independent measurement results obtained under specified boundary conditions. Repeatability conditions: 1 laboratory, 1 sample, 1 tester, 1 device, identical chemicals, short time intervals 1 laboratory, 1 sample, several testers, 1 device, measurement on different days Note: Measurement of different devices in the NTS within one task is not helpful due to the variability of the devices.	Reproducibility of the features (number, identity, signal intensity) with multiple sample measurements under the respective boundary conditions. Comparability of chromatography (long-term stability) The entire workflow (including peak finding) should be considered when testing the performance characteristic, as the validation includes the evaluation algorithms.	Carrying out the feature comparison of repeat measurements and determining the scatter RT, mass, and intensity (see Appendix J) Use of internal standards for each sample (see 7.5)
Robustness System stability	The ability of the method to be independent of minor fluctuations in the method parameters. - Chromatography - Mass spectrometer	Strength of the influence on the respective result according to the task. System stability of the mass spectrometer: - Mass stability - Resolving power	Deliberate variation of individual method parameters. Chromatography, e.g., temperature, gradient, modifier concentration Mass spectrometer: e.g., Source and interface temperature, gas flows

Feature	Description	Performance criteria Example	Implementation / Evaluation
		- Sensitivity These variables can be determined using control charts, for example, and evaluated using the device specifications. The control charts can be used across all methods. In routine use, they can be used as quality control charts (see chapter13.2)	Deliberate variation of several method parameters simultaneously: Design of Experiment (DoE) and Analytical Quality by Design (AQbD) approaches. Evaluation, e.g., by feature comparison (see Appendix J)

13 Quality assurance concept

13.1 Introduction

For the use of LC-ESI-HRMS(/MS) for NTS, the focus of quality assurance (QA) is on i) chromatographic separation performance, ii) ionization, and iii) mass spectrometric detection. The size of the data sets and the high information density of NTS data are a challenge for quality assurance compared to analytical methods of target analysis. By using the data as an archive for describing the condition of the test sample, quality assurance, and documentation are indispensable for long-term evaluations.

QA is crucial for NTS because:

- NTS data's size and information density make a manual plausibility check based on "expected" results difficult.
- NTS requires very stable measurements because "only" retention time, signal intensity, and correlating m/z values are often used to characterize a substance. Accordingly, these values must be recorded as correctly and precisely as possible and be verifiable.
- The strength of NTS measurements in environmental analysis often lies in comparing several samples (e.g., time series, before and after treatment processes or comparisons in the area) and in creating digital sample archives for retrospective data analysis. These applications require the smallest possible interday variance or sophisticated correction methods. For example, masses, retention times, and response require sufficient precision to map intensity changes of NT features in large data sets and thus perform statistical analyses.

A QA strategy should be based on several pillars that cover as many individual areas of the measurement system as possible. Several, sometimes duplicated, control points/control parameters are necessary to i) ensure redundancy/reliability, ii) cover as large an area as possible in the RT (polarity) and *m/z* domain, and iii) simplify troubleshooting, especially if only part of the system (a control point) is faulty.

13.2 QA parameters

13.2.1 General

Quality control charts have proven their worth for documenting and assessing QA parameters in analytics. Measurement stability is best controlled by quality control charts that are created, for example, by injecting (internal) standards at the beginning of a measurement sequence (LC+MS control), by direct infusion of standards (MS control only), or by combining both variants to obtain several control points (Figure 13-1). Possible parameters for the control charts are:

- Chromatographic retention time stability (optionally with *k*- or *α*- factor and signal half-width),
- Spectral or chromatographic signal intensity or corresponding peak area,
- Mass accuracy (comparison with the theoretical *m/z*) in ppm or mDa,
- Spectrometric resolution or spectral half-width.

As a rule, more than five standards (per MS level and polarity) are used to calculate the quality parameters mentioned. For this purpose, isotope-labeled internal standards are often measured in an ultrapure water sample. Ideally, these are distributed over the RT and *m/z* range to be measured. Quality control charts allow the retrospective evaluation of measurements' comparability at different times. Furthermore, the quality parameters can be viewed throughout individual sequences to observe drifts. Each manufacturer of LC-HRMS systems has its own control and analysis software. However, the use of quality control charts for the requirements of NTS is usually not (yet) integrated or must be extended, so individual solutions based on Excel, R, Matlab, etc. often have to be found (example in Appendix K).

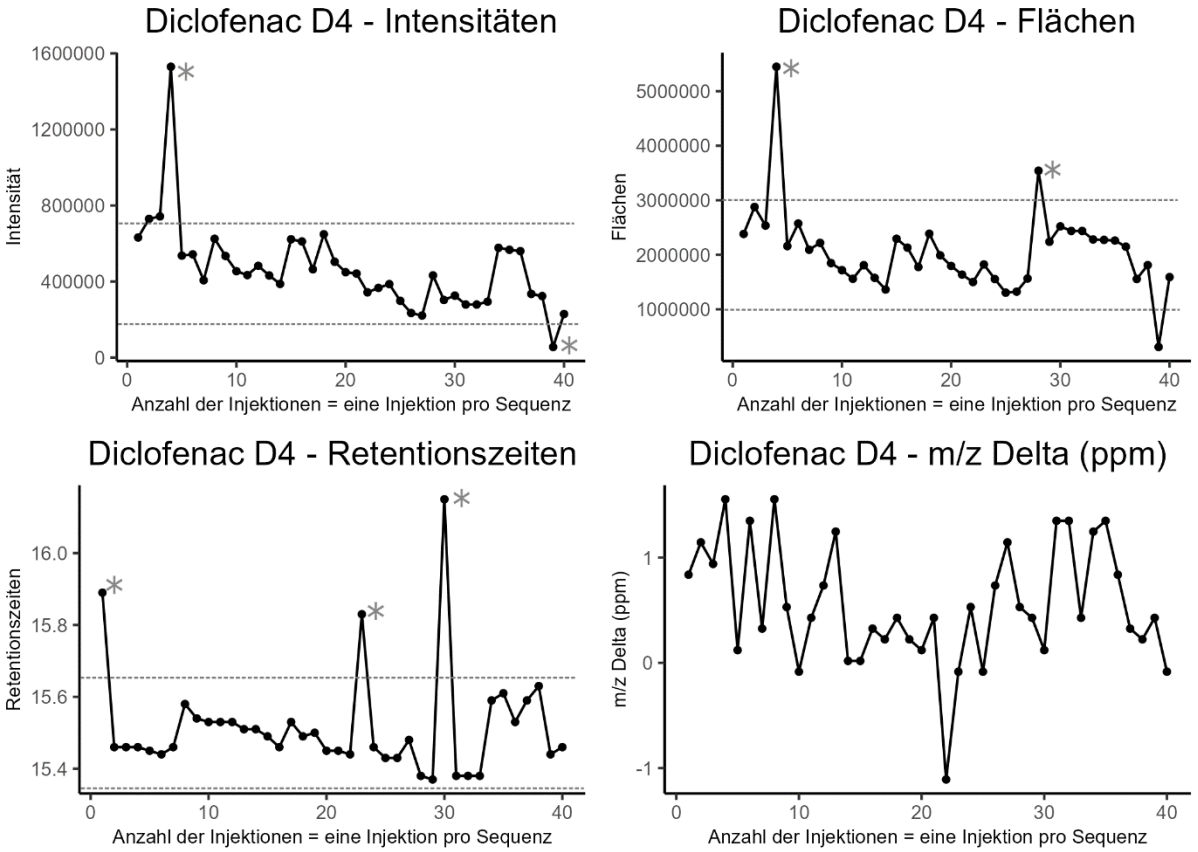


*Figure 13-1: Examples of some QA parameters over time. *: Out-of-control situation.*

13.2.2 Stability of chromatography

The retention time stability can be determined using appropriate reference substances. Typically, deviations of less than 0.1 min within a sequence and less than 0.3 min between sequences can be expected (Figure 13-1, bottom left). Changes in peak half-width and symmetry (tailing, fronting) also indicate the condition of the column or a possible matrix influence.

13.2.3 Intensity/sensitivity

The stability of the intensity/sensitivity is a combination of the stability of the chromatography, the ionization, and the mass spectrometer. The intensity is, therefore, a criterion for the robustness of the measurement and is thus crucial for a comparative evaluation of NTS data. For a QA of the routine's intensity (or peak area), the IS spiked to the samples can be used (see Figure 13-1, top row). Trends within a sequence can indicate increasing contamination of the ion source. A lower and upper device-specific warning limit for the intensity is recommended. These limits should be determined from a more extended previous period or a larger number

of measurements. This period should also include several calibrations of the mass spectrometer and cleanings of the ion source. A fluctuation in intensity or area by more than a factor of 2 to 3 (experience of FA members) should result in a search for the cause.

The experience of members of the FA shows that, for example, exceptionally high intensities fall back to stable "standard values" within a relatively short time after cleaning and calibration of the mass spectrometer. Therefore, a more extended observation period helps assess "exceptionally high" and "standard values," which are also regularly adjusted considering current measurements.

13.2.4 *m/z* accuracy and spectral resolution

For *m/z* accuracy and resolution, an upper or lower warning limit is sufficient. Depending on the instrumental setup, suitable warning limits can vary greatly. Experience has shown that an *m/z* accuracy greater than 5 ppm or above 5 mDa and a resolution of less than 20,000 is exceptional. If necessary, warning limits must be defined individually for both polarities and MS and MS^2 experiments.

13.2.5 Comparability

The use of isotope-labeled internal standards not only serves to monitor the QA parameters and the measurement system's stability but can also be used to compare different data sets. This applies to laboratory-internal comparability over more extended periods or measurement series recorded with different measurement systems and to data sets from different laboratories with different methods or devices. Uniform internal standards enable and improve the alignment of various data sets (see 10.1.3) through retention time correction and mass recalibration. For these purposes, a comprehensive mix of isotopically labeled internal standards should cover a broad mass and polarity range and contain as wide a variety of chemical functional groups as possible.

Based on the available experience of the Non-Target Screening Expert Committee members and application-specific criteria, the compilation of 22 substances shown in Table 13.1 is recommended as particularly suitable. To maximize the effect of better comparability, especially in inter-laboratory comparisons, including as many of the substances mentioned as possible in the NTS measurements is desirable.

To minimize effort and costs for users and to ensure good availability, the members of the FA NTS were able to convince a commercial supplier to produce and offer a ready-to-use mix. The mix can be obtained as a ready-to-use stock solution as "Pharma-Dx - Mix 22" under the article number " R01243-WaaAN5" from Neochema GmbH (Uwe-Zeidler-Ring 10, 55294 Bodenheim, Germany).

Specification ("Pharma-Dx - Mix 22"):

Concentration:	mostly 10 µg/ml, some compounds differ
Solvent:	Acetonitrile
Shelf life:	min. 12 months

Stock solutions with similar compositions from other manufacturers or self-produced mixtures of individual standards can be used similarly.

Table 13.1: Mix of isotope-labeled internal standards compiled from the experience of FA NTS members

Substance name	Sum formula	$[M+H]^+$	$[M-H]^-$	LogD (pH=5)	CAS No.
Metformin-D6	$C_4H_5D_6N_5$	J	N	-3.66	1185166-01-1
Sotalol-D7	$C_{12}D_7H_{13}N_2O_3S$	J	J	-3.18	1398065-65-0
Metoprolol-D7	$C_{15}H_{25}NO_3$	J	N	-1.47	959787-96-3
Acesulfame-D4	$C_4HD_4NO_4S$	N	J	-1.46	1623054-53-4
Hydrochlorothiazide-13C,D2	$^{13}C_1C_6H_6D_2ClN_3O_4S_2$	N	J	-0.58	1190006-03-1
Saccharin-D4	$C_7HD_4NO_3S$	N	J	-0.49	1189466-17-8
Iopromide-D3	$C_{18}H_{21}D_3I_3N_3O_8$	J	J	-0.44	1189947-73-6
Bentazon-D6	$C_{10}H_6D_6N_2O_3S$	N	J	-0.04	2733969-39-4
Sulfamethoxazole-D4	$C_{10}H_7D_4N_3O_3S$	J	J	0.76	1020719-86-1
Chloridazon-D5	$C_{10}H_3D_5ClN_3O$	J	J	1.11	1246818-99-4
Benzotriazole-D4	$C_6HN_3D_4$	J	J	1.30	1185072-03-0
Bromacil-D3	$C_9H_{10}D_3BrN_2O_2$	J	J	1.69	2714436-92-5
Simazine-D10	$C_7H_2D_{10}ClN_5$	J	N	1.77	220621-39-6
Diuron-D6	$C_9H_4D_6Cl_2N_2O$	J	J	1.93	1007536-67-5
DEET-D7	$C_{12}H_{10}D_7NO$	J	N	2.50	1219799-37-7
Carbamazepine-D10	$C_{15}H_2D_{10}N_2O$	J	N	2.77	132183-78-9
Bezafibrate-D6	$C_{19}H_{14}ClNO_4D_6$	J	J	2.79	1219802-74-0
Darunavir-D9	$C_{27}H_{28}D_9N_3O_7S$	J	J	2.82	1133378-37-6
Fipronil-^{13}C4	$C_8{}^{13}C_4H_4Cl_2F_6N_4OS$	J	J	4.49	2140327-54-2
Diflufenican-D3	$C_{19}H_8D_3F_5N_2O_2$	J	N	5.11	1185009-29-3
Irbesartan-D4	$C_{25}H_{24}D_4N_6O$	J	J	5.44	1216883-23-6
Telmisartan-D3	$C_{33}D_3H_{27}N_4O_2$	J	N	6.04	1189889-44-8

13.3 Sampling

As sampling errors significantly contribute to the overall error of the analytical process, quality assurance during sampling and transportation is of particular importance. Details can be found in DIN EN ISO 5667-14 *"Water quality - Sampling - Guidance for quality assurance and quality control in collecting and handling water samples".* The standard contains both preventive quality assurance measures and procedures for controlling the quality of water sampling.

Particular difficulties with NTS arise because the number and identity of the analytes and, therefore, their physico-chemical properties are generally not known a priori. Despite these difficulties, special attention must be paid to testing, minimizing, and avoiding contamination (keyword: blank values) and possible losses of analytes during sampling, transport, and storage.

DIN EN ISO 5667-14 proposes various quality control measures that are also applicable to NTS and are of fundamental importance:

- Various blank samples for testing for contamination (environmental blank sample, field blank sample, transport blank sample, equipment blank sample, filter blank sample)
- Duplicate samples to calculate the precision of the sampling
- Doping with standards to test the stability of the sample

A possible procedure for taking a groundwater sample with a submersible pump is shown as an example. If possible, sampling with ultrapure water should be "simulated" in the field to obtain blank values. For this purpose, the sampling equipment (submersible pump, hoses, vessels with lids) that comes into contact with the sample water is included, and thus, a sample is generated that allows the blank values of the sampling to be determined.

Doped ultrapure water can be used as a sample to check for any losses during sampling. Ideally, the selection of substances should cover the entire range of physico-chemical properties of the non-target analytes (e.g., molar mass, K_{ow}, *logD*, functional groups, charge) and thus be representative of the analyte spectrum recorded in the NTS. The selection of substances can also depend on the task (e.g., searching for pharmaceuticals or pesticides).

Potential additional contamination and loss of analytes during sample transportation and storage can be controlled with field blanks. These are spiked ultrapure water samples taken into the field in the sampling containers and treated the same way as real samples. For example, a field blank sample can remain sealed and thus describe any changes in the sample due to transportation compared to a doped subsample that remained in the laboratory. Individual steps in manipulating samples, such as filtration, decanting, and dividing, can be described similarly.

Another component of QA is the spiking of an environmental sample with isotope-labeled standards, which can be used to detect sample-specific changes due to precipitation, biodegradation, sorption on or desorption of particles, etc.

13.4 Sample preparation

Sample preparation can significantly contribute to the analytical process's overall error. Concerning quality assurance during sample preparation, the following aspects should be considered:

- Checking the sample preparation for losses or contamination
    - By adding standards at the beginning of the sample preparation and comparatively in ready-to-measure undoped samples immediately before the measurement.
    - Comparison of the NTS result of a sample before and after sample preparation [5]
- Measurement of field blank values (consideration of blank values from the sampling materials)
- Measurement of the method blank values over the entire process: Preparation of extraction blanks that undergo all preparation steps (filtration, SPE, etc.), including using the same vessels and contact with the same materials.

14 Comparative study

14.1 Participants

Name	Institution / (for the trial period in the years 2014/2015)
Brüggen, Susanne	State Agency for Nature, Environment, and Consumer Protection NRW D - 47051 Duisburg
Dünnbier, Uwe	Laboratory of Berliner Wasserbetriebe (BWB) D - 13629 Berlin
Fink, Angelika Götz, Sven	Hessenwasser GmbH & Co KG D - 64293 Darmstadt
Geiß, Sabine	Thuringian State Institute for Environment and Geology Environmental analysis/environmental radioactivity D-07745 Jena
Letzel, Thomas Grosse, Sylvia	Technical University of Munich (TUM) AFG, Chair of Urban Water Management D - 80333 Munich
Petri, Michael	Zweckverband Bodensee-Wasserversorgung D - 78354 Sipplingen
Scheurer, Marco	DVGW Water Technology Center D - 76139 Karlsruhe
Schlüsener, Michael Kunkel, Uwe	Federal Institute of Hydrology D - 56068 Koblenz
Schulz, Wolfgang Lucke, Thomas	Zweckverband Landeswasserversorgung D - 89129 Langenau
Singer, Heinz	Eawag CH - 8600 Dübendorf
Stötzer, Sebastian	Bachema AG CH - 8952 Schlieren
Schlett, Claus	Westfälische Wasser- und Umweltanalytik GmbH D - 45891 Gelsenkirchen
Seiwert, Bettina	Department Analytics Helmholtz Center for Environmental Research GmbH - UFZ D - 04318 Leipzig
Sengl, Manfred	Bavarian State Office for the Environment D - 86179 Augsburg
Türk, Jochen	Institute for Energy and Environmental Technology e.V. (IUTA D - 47229 Duisburg
Zwiener, Christian	University of Tübingen Environmental analysis at ZAG D - 72074 Tübingen

The comparative studies were partly carried out in the BMBF project FOR-IDENT.

14.2 Implementation

Two comparative measurements were carried out in 2014 and 2015 as part of the "Non-Target Screening" technical committee of the Water Chemistry Society (see chapter14.1

14.2.1 Comparative measurement A

- Participants:
  - Sent to 18 participants (response rate 15 data sets)
  - MS manufacturers: Agilent, SCIEX, Thermo, Waters
- Sample set:
  - Blanks and methanolic reference standards (10 mg/L) for dilution by the participant
  - 5 substances each for positive and negative electrospray ionization
  - 2 additional substances with ionizability in both ESI modes
- Specifications:
  - A fixed injection volume of 10 µL (comparative evaluation of device sensitivity)
  - Literature spectra of the known compounds
- Analytics:
  - (Suspect) target screening for the known compounds using the LC-HRMS methods established with the participants
- Task:
  - Dilution of the standard solution in decadic steps
  - Single measurement of the dilutions to determine the detection limits (detection of at least two of the specified fragment ions)
  - Comparison of production spectra with literature spectra
  - Triple measurement at the detection limit
- Recorded data:
  - Applied methodology
  - Precursor masses
  - Detection limits

14.2.2 Comparative measurement B

- Participants:
  - 21 participants (response rate 18 data sets)
  - MS manufacturers: Agilent, SCIEX, Bruker, Thermo, Waters
- Sample set:
  - 4 randomly spiked Danube water samples (undoped, 0.025, 0.10 and 0.50 µg/L)
  - 24 doped compounds (not known to the participant but included in the suspect list)
- Specifications:
  - Suspect/non-target screening (established workflow in each case)
  - Suspect list (approx. 200 substances)
  - RTI-Std. (TUM) - Data return and evaluation TUM
- Analytics:
  - Established screening workflow (suspect or non-target)
- Task:
  - Determination of the doped compounds
  - Testing the sum formula (isotope)
  - Type of identification (database, reference standard)
  - Identification and categorization (according to)0

14.3 Results

14.3.1 Methods used

The separation was performed on a reversed-phase chromatography using methanol or acetonitrile with an ionization auxiliary in the acidic pH range for all participants. All participants used electrospray ionization in both positive and negative measurement modes. The automated acquisition of MS/MS spectra in the same run depended on the mass spectrometer data acquisition speed. If automated acquisition was impossible, MS/MS spectra were acquired in separate runs and used for analysis.

14.3.2 Sensitivity

The sensitivity of the systems was evaluated by diluting the methanolic solutions of 10 mg/L per substance in steps of 10 with water. The detection limit was the dilution at which two of the specified fragment ions could still be detected at 10 µL injection volume (Figure 14-1).

		Detektionsgrenzen in µg/L										
		Nr. 1	Nr. 6	Nr. 7	Nr. 8	Nr. 10	Nr. 11	Nr. 13	Nr. 2	Nr. 3	Nr. 4	Nr. 12
ESI pos	Alachlor	1	100	1	1	100	10	0.1	0.01	100	0.1	0.01
	Atrazin	0.1	10	0.1	0.1	10	0.1	0.1	0.1	0.0001	0.01	0.01
	Clarithromycin	0.1	1000	1	1	10	1	0.1	0.1	0.001	0.1	0.01
	Gabapentin	1	100	0.1	1	10	1	0.1	n.n	0.0001	0.01	1
	Quinoxyfen	0.1	1000	0.1	0.1	10	0.1	0.1	0.01	0.0001	0.1	0.01
	Valsartan	0.1	n.n.	0.1	1	100	1	0.1	0.01	0.0001	0.01	1
	Candesartan	0.001	100	0.1	0.1	1	1	0.1	0.1	0.0001	0.01	1
ESI neg	PFNA	n.n.	1000	0.1	1	100	10	n.a.	1	0.0001	0.1	0.1
	HCT	1	1000	1	1	10	10	n.a.	0.1	0.0001	0.1	1
	Mecoprop	1	1000	1	1	10	10	n.a.	0.1	0.0001	0.1	1
	Ioxynil	0.01	1000	0.1	0.1	1	1	n.a.	0.01	0.0001	0.1	0.1
	Dinoseb	0.01	100	0.1	0.1	0.1	0.1	n.a.	0.01	0.0001	0.1	0.01
	Valsartan	0.01	n.n.	0.1	1	10	10	n.a.	0.1	0.0001	0.1	1
	Candesartan	0.001	100	0.1	1	10	10	n.a.	0.1	0.001	0.1	1
nicht vermessen/analysiert												

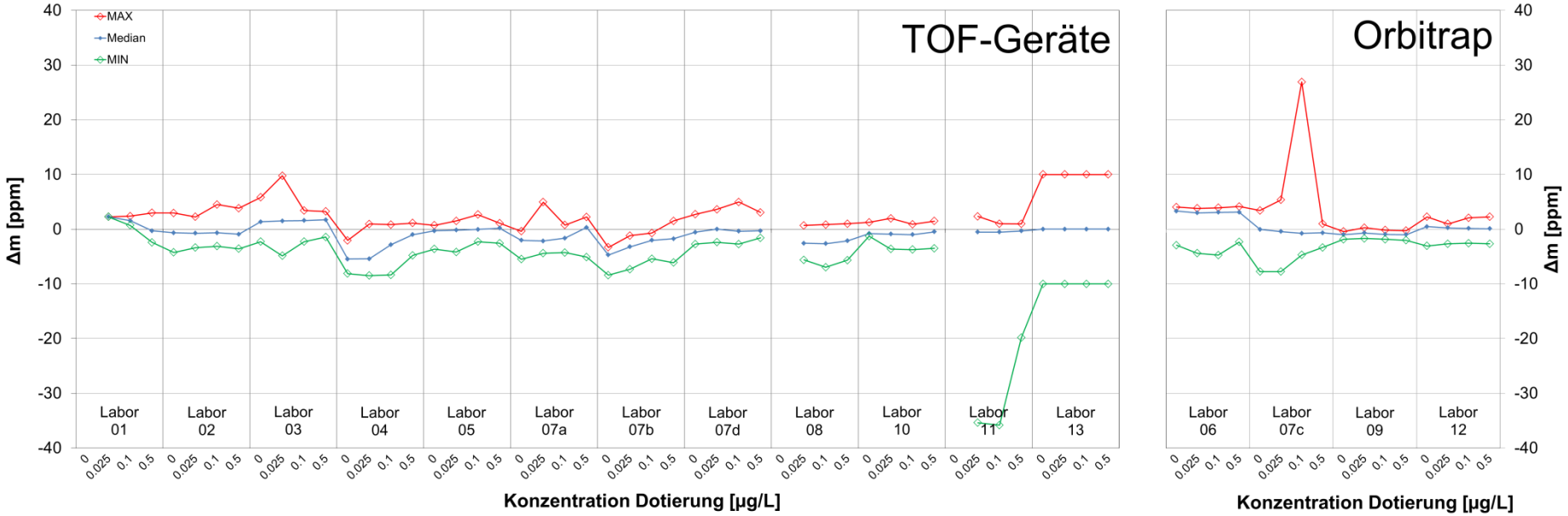
TOF-Geräte OrbiTrap

Figure 14-1: Comparison of the detection limits as concentration with at least two detectable fragment ions (labs 6 and 3 outliers), PFNA: perfluorononanoic acid, HCT: hydrochlorothiazide

14.3.3 Mass accuracy MS

The median of the mass deviations of the molecular ions of the doped compounds was less than 5 ppm. There were no differences in the mass accuracy of the different TOF and Orbitrap systems. The mass deviations were also independent of the doped concentration (Figure 14-2).

Figure 14-2: Mass deviations in MS mode (laboratories 8 and 11: undoped sample not measured)



14.3.4 Mass accuracy of the fragment masses (MS/MS)

Qualitative differences in the fragment ion spectra were mainly due to the collision energy. If the mass accuracy of the fragments is considered, there is a difference between the TOF and Orbitrap devices. Time-of-flight mass spectrometers (Figure 14-3) show a slightly larger mass deviation in MS/MS experiments compared to Orbitrap devices (Figure 14-4). The deviations for TOF devices are usually less than 5 mDa, corresponding to a relative deviation of 5 to 50 ppm. For Orbitrap devices, the absolute mass deviations are usually below 2 mDa, corresponding to a relative deviation of 2 to 40 ppm (mass range *m/z* 50 - 1000).

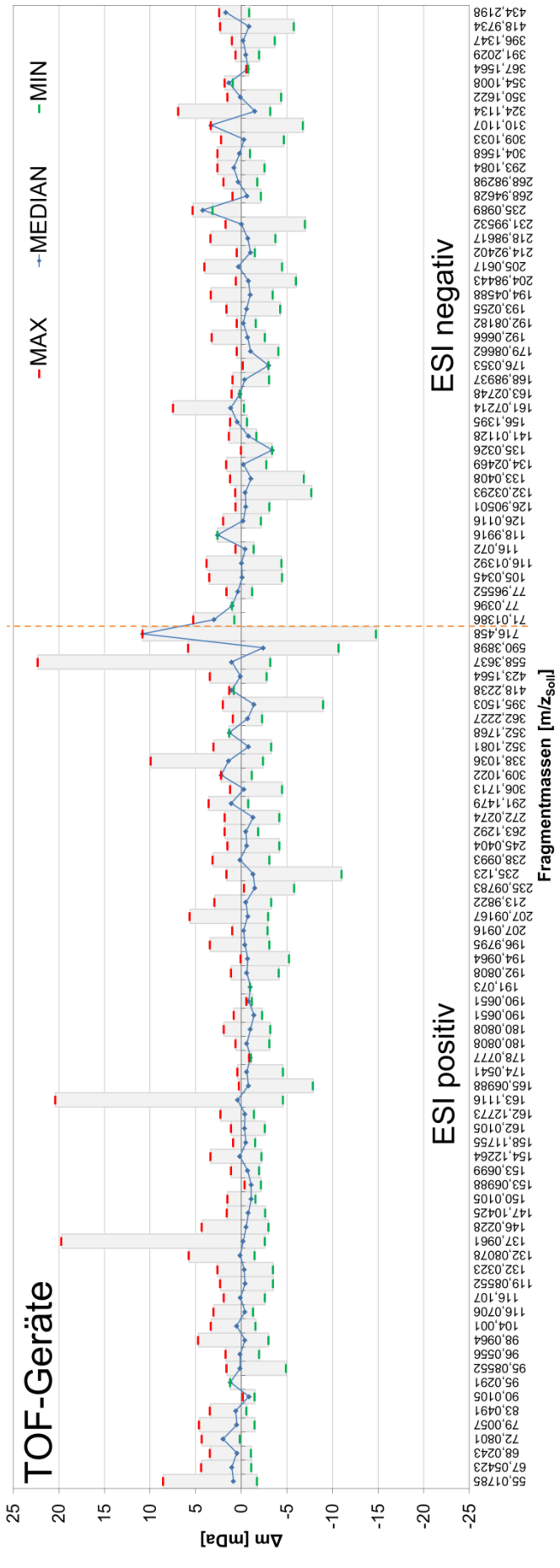
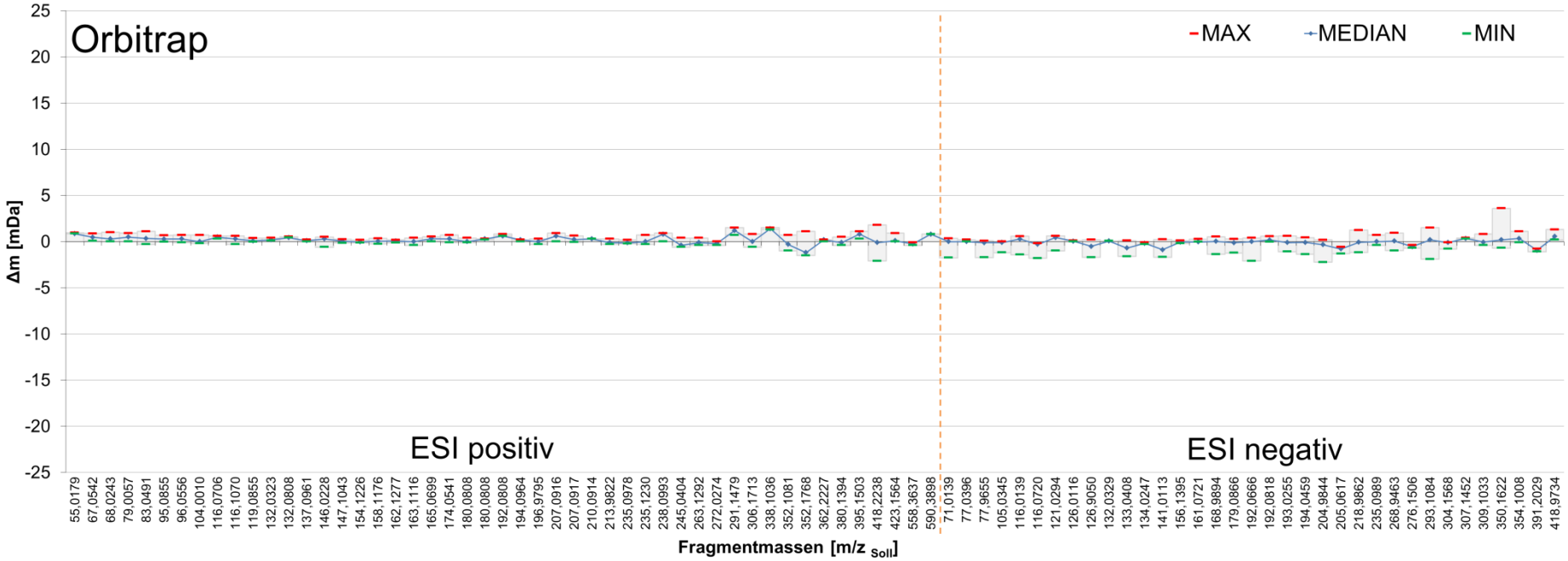


Figure 14-3: Mass deviations of the MS/MS fragments of the doped compounds (TOF devices); ordered by fragment mass and separated by ionization mode

Figure 14-4: *Mass deviations of the MS/MS fragments of the doped compounds (Orbitrap devices); ordered by fragment mass and separated by ionization mode*

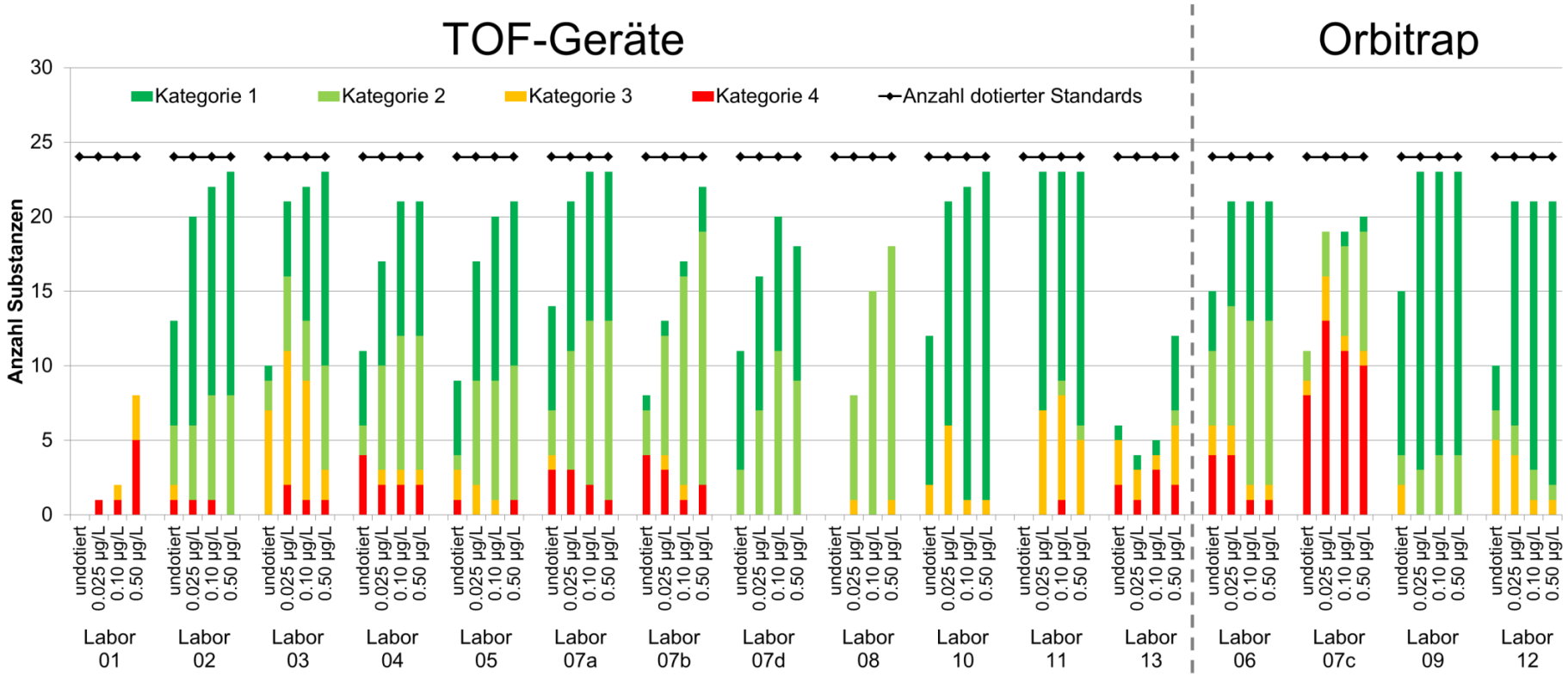


14.3.5 Data evaluation and substance identification

The numbers of correctly identified standard substances in the participating laboratories are shown in Figure 14-5. The detected signals were categorized according to the criteria described in chapter 10.2.1. The increase in the proportion of representatives in Category 1 (reliable substance identification) and Category 2 (hits with high probability) with increasing doping levels is clearly recognizable. This is generally due to the improved detectability of a meaningful MS/MS spectrum.

Laboratory 7 is a special case. The participation of a laboratory with a total of four LC-HRMS systems (a to d, each with a specific operator) makes it clear that the system used (in particular the software options) and the available database (measured reference standards and MS^2 spectra) have a significant influence on the number of identifications. Laboratory 7c shows a significantly lower number of correctly identified substances. The number of qualitative detections was comparable to other systems. This may be due to low available reference spectra or a complex software solution for the identification step. Last but not least, the experience of the user and the time required for data evaluation also play a decisive role.

Figure 14-5: Comparison of the standard substances identified by the participating laboratories according to the identification categories 1 to 4



14.3.6 Workflow comparison using the example of a laboratory

In addition to the comparative measurement, one of the data sets from the second comparative measurement was evaluated using three different workflows to examine the influence of the approach on the number of correctly identified substances (Figure 14-6).

The three workflows used were structured as follows:

(1) **Suspect** screening for the entire suspect list (200 compounds) and their manual evaluation concerning identity with a comparison of MS^2-spectrum libraries

(2) **Non-target approach with peak finding using the open-source tool envipy**[1] and subsequent manual evaluation of the identification against reference spectra (the envipy tool is no longer available. The successor tool enviMass is no longer open-source).

(3) A **non-target approach (internal laboratory)** with data evaluation and subsequent FOR-IDENT query will prioritize the possible substance proposals for the signals. Identification utilizing reference MS^2spectral database.

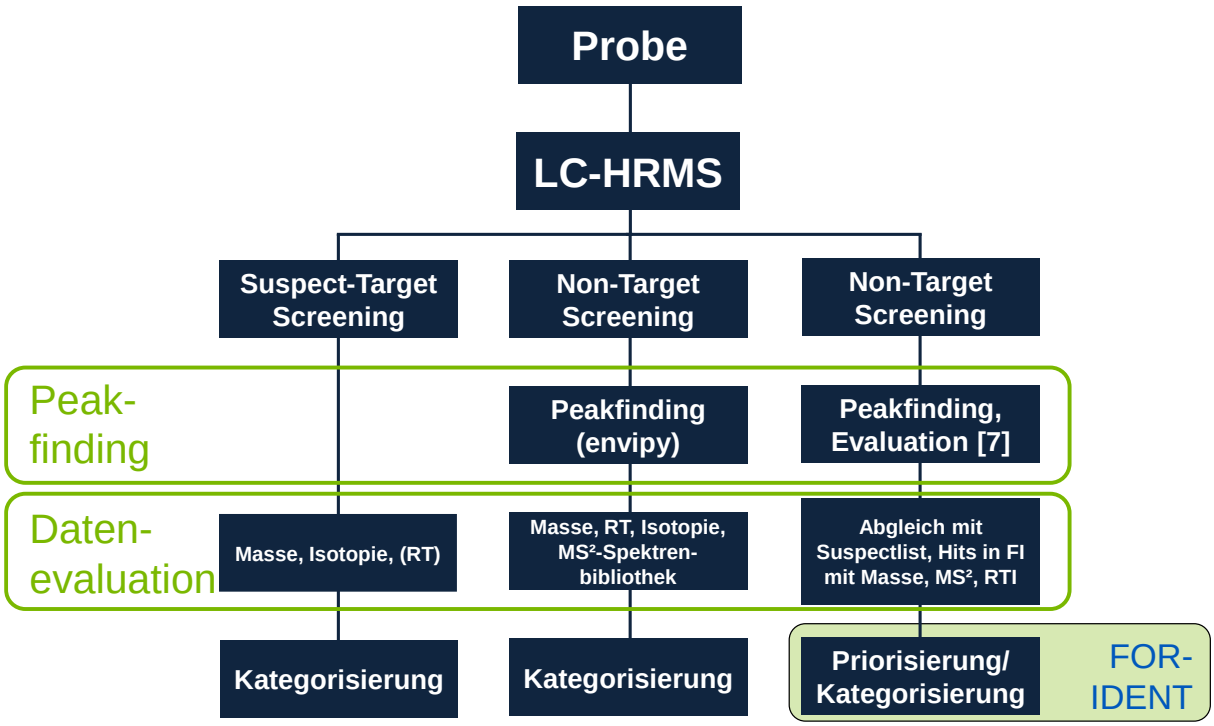


Figure 14-6: Structure of three different workflows (1), (2) and (3) for the detection and identification of substances

[1] https://www.envibee.ch/eng/projects.htm

A comparison of the results of the three different workflows (1), (2) and (3) (Figure 14-7) shows good detectability of the doped compounds. In case 2 (Figure 14-7, center), the number of detected compounds (categories 1 to 4) is slightly below the other two workflows. This could be due to insufficient optimization of the peak finding parameters. The peak finding in the third workflow was developed on the LC-HRMS system used for the measurement and is, therefore, certainly best suited for this system. This is reflected in the highest detection figures. The possibility of identification (MS^2 spectra, databases) was the same in all cases, which is reflected in a barely different number of substances in categories 1 and 2. Therefore, automation's advantages are most evident in the time required. The detection of the compounds searched for was hardly influenced by the choice of workflow.

The first workflow (suspect-target screening) was the most time-consuming, requiring processing and manual review of the 200 substances for identification. In addition, all reference spectra not contained in the existing spectra library had to be researched and compared in Internet databases. However, the manual effort of the three workflows decreases significantly from (1) to (3). This is partly due to the automated peak finding in cases (2) and (3), but also in particular to the automated prioritization of the substance proposals by FOR-IDENT in case (3). As expected, the number of doped compounds detected increases with the concentration of the substances

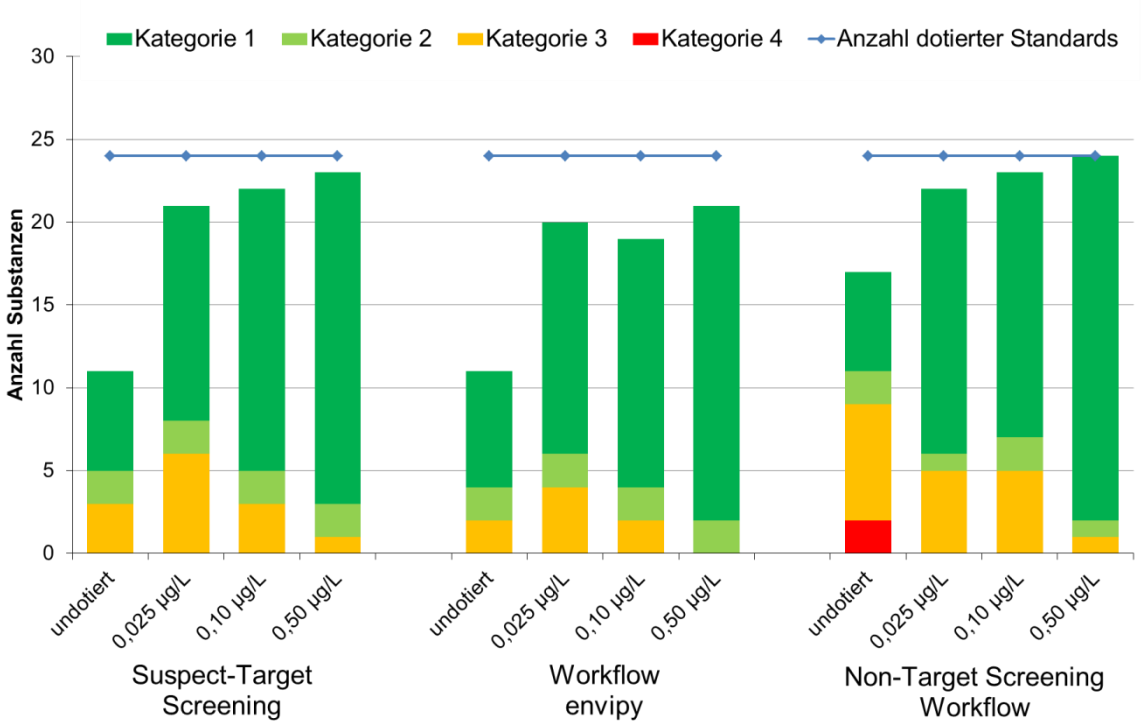


Figure 14-7: Comparison of identification results of a data set with three different evaluation workflows (1), (2), (3)

15 Bibliography

1. Hollender J, Schymanski EL, Singer HP, Ferguson PL. Nontarget Screening with High Resolution Mass Spectrometry in the Environment: Ready to Go? Environ Sci Technol. 2017;51(20):11505-12.

2. Nürenberg G, Schulz M, Kunkel U, Ternes TA. 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. Journal of Chromatography A. 2015;1426:77-90.

3. Bader T, Schulz W, Lucke T, Seitz W, Winzenbacher R. Application of Non-Target Analysis with LC-HRMS for the Monitoring of Raw and Potable Water: Strategy and Results. Assessing Transformation Products of Chemicals by Non-Target and Suspect Screening − Strategies and Workflows Volume 2. ACS Symposium Series. 1242: American Chemical Society; 2016. p. 49-70.

4. Alygizakis NA, Samanipour S, Hollender J, Ibáñez M, Kaserzon S, Kokkali V, et al. Exploring the Potential of a Global Emerging Contaminant Early Warning Network through the Use of Retrospective Suspect Screening with High-Resolution Mass Spectrometry. Environ Sci Technol. 2018;52(9):5135-44.

5. Bader T, Schulz W, Kümmerer K, Winzenbacher R. LC-HRMS data processing strategy for reliable sample comparison exemplified by the assessment of water treatment processes. Analytical chemistry. 2017;89(24):13219-26.

6. Hollender J, Schymanski EL, Ahrens L, Alygizakis N, Béen F, Bijlsma L, et al. NORMAN guidance on suspect and non-target screening in environmental monitoring. Environmental Sciences Europe. 2023;35(1):75.

7. Vosough M, Schmidt TC, Renner G. Non-target screening in water analysis: recent trends of data evaluation, quality assurance, and their future perspectives. Analytical and Bioanalytical Chemistry. 2024:1-12.

8. Murray KK, Boyd RK, Eberlin MN, Langley GJ, Li L, Naito Y. Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013). Pure and Applied Chemistry. 2013;85(7):1515-609.

9. Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Molecular & Cellular Proteomics. 2012;11(6).

10. Bader T, Schulz W, Kümmerer K, Winzenbacher R. General strategies to increase the repeatability in non-target screening by liquid chromatography-high resolution mass spectrometry. Analytica Chimica Acta. 2016;935:173-86.

11. Kind T, Fiehn O. Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. BMC bioinformatics. 2007;8(1):1-20.

12. MassBank Europe [Internet]. Available from: https://massbank.eu/MassBank/.

13. mzCloud [Internet]. Available from: https://www.mzcloud.org/.

14. PubChem [Internet]. Available from: https://pubchem.ncbi.nlm.nih.gov/.

15. ChemSpider [Internet]. Available from: http://www.chemspider.com/.

16. Wolf S, Schmidt S, Müller-Hannemann M, Neumann S. In silico fragmentation for computer assisted identification of metabolite mass spectra. BMC bioinformatics. 2010;11:1-12.

17. STOFF-IDENT Datensatz [Internet]. Available from: https://water.for-ident.org/#!home.

18. FOR-IDENT [Internet]. Available from: http://www.chemspider.com/.

19. CompTox Chemicals Dashboard [Internet]. Available from: https://www.epa.gov/comptox-tools/comptox-chemicals-dashboard.

20. DIN, ISO. DIN ISO 5667-5:2011-02 : Wasserbeschaffenheit - Probenahme - Teil 5: Anleitung zur Probenahme von Trinkwasser aus Aufbereitungsanlagen und Rohrnetzsystemen. 2011. p. 25.

21. Schulze B, Heffernan AL, Thomas KV, Kaserzon SL. Influence of Sample Stability on Non-target Analysis during Long Batch Acquisitions. ACS ES&T Water. 2023;3(9):2874-82.

22. DIN, ISO. DIN EN ISO 5667-6:2016-12: Wasserbeschaffenheit - Probenahme - Teil 6: Anleitung zur Probenahme aus Fließgewässern. 2016. p. 41.

23. DIN. 38402-11:2009-02 Deutsche Einheitsverfahren zur Wasser-, Abwasser- und Schlammuntersuchung - Allgemeine Angaben (Gruppe A) - Teil 11: Probenahme von Abwasser. 2009. p. 21.

24. Aalizadeh R, Alygizakis NA, Schymanski EL, Krauss M, Schulze T, Ibanez M, et al. Development and application of liquid chromatographic retention time indices in HRMS-based suspect and nontarget screening. Analytical chemistry. 2021;93(33):11601-11.

25. 2002/657/EG: Entscheidung der Kommission vom 12. August 2002 zur Umsetzung der Richtlinie 96/23/EG des Rates betreffend die Durchführung von Analysemethoden und die Auswertung von Ergebnissen, (2002).

26. ISO. 21253-1: 2019: Water quality - Multi-compound class methods – Part 1: Criteria for the identification of target compounds by gas and liquid chromatography and mass spectrometry. In: 2 ITS, editor. 1 ed. 21253-1: 20192019. p. 21.

27. DIN. DIN 38407-47:2017-07: Bestimmung ausgewählter Arzneimittelwirkstoffe und weiterer organischer Stoffe in Wasser und Abwasser - Verfahren mittels Hochleistungs-Flüssigkeitschromatographie und massenspektrometrischer Detektion (HPLC-MS/MS oder -HRMS) nach Direktinjektion. In: 47; F, editor. 2017. p. 42.

28. Evaluation of Analytical Methods for EDCs and PPCPs via Inter-Laboratory Comparison [Internet]. 2012-2023. Available from: https://www.waterrf.org/research/projects/evaluation-analytical-methods-edcs-and-ppcps-inter-laboratory-comparison.

29. Raetz M, Bonner R, Hopfgartner G. SWATH-MS for metabolomics and lipidomics: critical aspects of qualitative and quantitative analysis. Metabolomics. 2020;16:1-14.

30. Helmus R, ter Laak TL, van Wezel AP, de Voogt P, Schymanski EL. patRoon: open source software platform for environmental mass spectrometry based non-target screening. Journal of Cheminformatics. 2021;13(1):1.

31. Schmid R, Heuckeroth S, Korf A, Smirnov A, Myers O, Dyrlund TS, et al. Integrative analysis of multimodal mass spectrometry data in MZmine 3. Nature biotechnology. 2023;41(4):447-9.

32. enviMass - mass spec analysis workflow [Internet]. Available from: https://www.envibee.ch/eng/enviMass/overview.htm.

33. Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Analytical chemistry. 2006;78(3):779-87.

34. Schollée JE, Schymanski EL, Hollender J. Statistical approaches for LC-HRMS data to characterize, prioritize, and identify transformation products from water treatment processes. Assessing Transformation Products of Chemicals by Non-Target and Suspect Screening− Strategies and Workflows Volume 1. 2016:45-65.

35. Katajamaa M, Orešič M. Data processing for mass spectrometry-based metabolomics. Journal of chromatography A. 2007;1158(1-2):318-28.

36. Myers OD, Sumner SJ, Li S, Barnes S, Du X. Detailed investigation and comparison of the XCMS and MZmine 2 chromatogram construction and chromatographic peak detection methods for preprocessing mass spectrometry metabolomics data. Analytical Chemistry. 2017;89(17):8689-95.

37. Tisler S, Pattison DI, Christensen JH. Correction of matrix effects for reliable non-target screening LC–ESI–MS analysis of wastewater. Analytical Chemistry. 2021;93(24):8432-41.

38. Kuhl C, Tautenhahn R, Bottcher C, Larson TR, Neumann S. CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. Analytical chemistry. 2012;84(1):283-9.

39. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, et al. Identifying small molecules via high resolution mass spectrometry: communicating confidence. ACS Publications; 2014.

40. Tejada-Casado C, Hernández-Mesa M, Monteau F, Lara FJ, del Olmo-Iruela M, García-Campaña AM, et al. Collision cross section (CCS) as a complementary parameter to characterize human and veterinary drugs. Analytica chimica acta. 2018;1043:52-63.

41. Horai H, Arita M, Kanaya S, Nihei Y, Ikeda T, Suwa K, et al. MassBank: a public repository for sharing mass spectral data for life sciences. Journal of Mass Spectrometry. 2010;45(7):703-14.

42. Slobodnik J, Hollender J, Schulze T, Schymanski EL, Brack W. Establish data infrastructure to compile and exchange environmental screening data on a European scale. Environmental Sciences Europe. 2019;31(1):65.

43. Sheldon MT, Mistrik R, Croley TR. Determination of ion structures in structurally related compounds using precursor ion fingerprinting. Journal of the American Society for Mass Spectrometry. 2011;20(3):370-6.

44. Aalizadeh R, Nika M-C, Thomaidis NS. Development and application of retention time prediction models in the suspect and non-target screening of emerging contaminants. Journal of Hazardous materials. 2019;363:277-85.

45. NORMAN Network [Internet]. Available from: https://www.norman-network.net/.

46. Samanipour S, Reid MJ, Thomas KV. Statistical variable selection: an alternative prioritization strategy during the nontarget analysis of LC-HR-MS data. Analytical chemistry. 2017;89(10):5585-91.

47. Bertinetto C, Engel J, Jansen J. ANOVA simultaneous component analysis: A tutorial review. Analytica Chimica Acta: X. 2020;6:100061.

48. Tautenhahn R, Böttcher C, Neumann S. Highly sensitive feature detection for high resolution LC/MS. BMC bioinformatics. 2008;9(1):1-16.

49. Vergeynst L, Van Langenhove H, Demeestere K. Balancing the false negative and positive rates in suspect screening with high-resolution Orbitrap mass spectrometry using multivariate statistics. Analytical chemistry. 2015;87(4):2170-7.

50. Westad F, Marini F. Validation of chemometric models–a tutorial. Analytica Chimica Acta. 2015;893:14-24.

51. DIN, ISO, EN. DIN EN ISO/IEC 17025:2018-03: Allgemeine Anforderungen an die Kompetenz von Prüf- und Kalibrierlaboratorien (ISO/IEC 17025:2017). 2018. p. 65.

52. Uhlig S, Nichani K, Stoyke M, Gowik P. Validation of binary non-targeted approaches: mathematical framework and experimental designs. bioRxiv. 2021:2021.01. 19.427235.

53. Fisher CM, Peter KT, Newton SR, Schaub AJ, Sobus JR. Approaches for assessing performance of high-resolution mass spectrometry–based non-targeted analysis methods. Analytical and Bioanalytical Chemistry. 2022;414(22):6455-71.

54. Kruve A. Strategies for Drawing Quantitative Conclusions from Nontargeted Liquid Chromatography–High-Resolution Mass Spectrometry Analysis. Analytical Chemistry. 2020;92(7):4691-9.

55. Keller BO, Sui J, Young AB, Whittal RM. Interferences and contaminants encountered in modern mass spectrometry. Analytica Chimica Acta. 2008 Oktober:71-81.

Appendix A. “Non-target screening" expert committee

A.1 Background and tasks

The Non-Target Screening Technical Committee was founded in 2009 in the Water Chemistry Society (a specialist group within the German Chemical Society). The starting point was to support the identification of trace substances in analysis using LC-MS by setting up a suitable database (also usable when using low-resolution systems). With the development of high-resolution mass spectrometers for routine use, the focus has shifted towards target analysis, suspect-target, and non-target screening. In future work, the group will focus on the following topics:

- Tools for identification
- Cross-laboratory NTS
- Quality assurance during data processing

A.2 Members of the FA (as of 2024)

Table A.1: Members of the "Non-Target Screening" expert committee

Name	Institution/Address
Management:	
Bader, Tobias and Jewell, Kevin	Zweckverband Landewasserversorgung Am Spitzigen Berg 1 D-89129 Langenau Federal Institute of Hydrology Am Mainzer Tor 1 D-56068 Koblenz
Members:	
Armbruster, Dominic	DVGW Water Technology Center Karlsruher Strasse 84 D-76139 Karlsruhe
Brüggen, Susanne	State Office for Nature, Environment and Consumer Protection NRW Duisburg office Wuhanstrasse 6 D-47051 Duisburg
Cunha, Ricardo	Institute for Energy and Environmental Technology e.V. (IUTA) Bliersheimer Str. 58 - 60 D-47229 Duisburg
Goetz, Sven	Hessenwasser GmbH & Co KG Gräfenhäuser Strasse 118 D-64293 Darmstadt
Härtel, Christoph	Ruhrverband Kronprinzenstrasse 37 D-45128 Essen
Käberich, Merle	DVGW Water Technology Center Karlsruher Strasse 84 D-76139 Karlsruhe
Kronsbein, Anna Lena	Federal Environment Agency Colditzstrasse 34 D-12099 Berlin
Kunkel, Uwe	Bavarian State Office for the Environment Bürgermeister-Ulrich-Strasse 160 D-86179 Augsburg
Letzel, Thomas	Analytical Research Institute for Non-Target Screening Am Mittleren Moos 48 D-86167 Augsburg

Name	Institution/Address
Liebmann, Diana	Laboratory of Berliner Wasserbetriebe (BWB) Motardstrasse 35 D-13629 Berlin
Liesener, André	Westfälische Wasser- und Umweltanalytik GmbH Willy-Brandt-Allee 26 D-45891 Gelsenkirchen
Logemann, Jörg	Free and Hanseatic City of Hamburg, Authority for Health and Consumer Protection, Institute for Hygiene and Environment Marckmannstrasse 129b D-20539 Hamburg
Lucke, Thomas	Zweckverband Landeswasserversorgung Am Spitzigen Berg 1 D-89129 Langenau
Macherius, André	Bavarian State Office for the Environment Bürgermeister-Ulrich-Strasse 160 D-86179 Augsburg
Merkus, Valentina	Hamburger Wasserwerke GmbH Billhorner Deich 2 D-20539 Hamburg
Petri, Michael	Zweckverband Bodensee-Wasserversorgung Süssenmühle 1 D-78354 Sipplingen
Reineke, Anna	Westfälische Wasser- und Umweltanalytik GmbH Willy-Brandt-Allee 26 D-45891 Gelsenkirchen
Renner, Gerrit	University of Duisburg-Essen, Instrumental Analytical Chemistry (IAC) Universitätsstrasse 5 D-45141 Essen
Ruppe, Steffen	Department of Economic, Social and Environmental Environmental Laboratory Department Spiegelgasse 15 CH-4001 Basel
Scheurer, Marco	LUBW State Institute for the Environment Baden-Württemberg Griesbachstrasse 1 D-76185 Karlsruhe
Schlüsener, Michael	Federal Institute of Hydrology Am Mainzer Tor 1 D-56068 Koblenz
Schulz, Wolfgang	Aalen University - Technology and Economics Beethovenstrasse 1 D-73430 Aalen
Singer, Heinz	Eawag-Swiss Federal Institute of Aquatic Science and Technology Ueberlandstrasse 133 CH-8600 Dübendorf
Türk, Jochen	Cooperation laboratory of Ruhrverband, Emschergenossenschaft and Lippeverband Kronprinzenstrasse 37 D-45128 Essen
Zahn, Daniel	Helmholtz Center for Environmental Research Permoserstrasse 15 D-04318 Leipzig
Zwiener, Christian	University of Tübingen, Environmental analysis at ZAG Hölderlinstrasse 121 D-72074 Tübingen

Appendix B. Mass and RT control

B.1 Isotope-labeled internal standards

Table B.1: Mix of isotope-labeled internal standards compiled from the experience of FA NTS members

Substance	Sum formula	Neutral mass M	$[M+H]^+$	$[M-H]^-$	LogD (pH=5)	CAS
Metformin-D6	$C_4H_5D_6N_5$	135.1391	J	N	-3.66	1185166-01-1
Sotalol D7	$C_{12}D_7H_{13}N_2O_3S$	279.1634	J	J	-3.18	1398065-65-0
Metoprolol-d7	$C_{15}H_{25}NO_3$	274.2274	J	N	-1.47	959787-96-3
Acesulfame-D4	$C_4HD_4NO_4S$	167.0190	N	J	-1.46	1623054-53-4
Hydrochlorothiazide 13C,D2	$^{13}C_1C_6H_6D_2ClN_3O_4S_2$	299.9804	N	J	-0.58	1190006-03-1
Saccharin D4	$C_7HD_4NO_3S$	187.0241	N	J	-0.49	1189466-17-8
Iopromide-D3	$C_{18}H_{21}D_3I_3N_3O_8$	793.8886	J	J	-0.44	1189947-73-6
Bentazon D6	$C_{10}H_6D_6N_2O_3S$	246.0945	N	J	-0.04	2733969-39-4
Sulfamethoxazole-D4	$C_{10}H_7D_4N_3O_3S$	257.0772	J	J	0.76	1020719-86-1
Chloridazon-D5	$C_{10}H_3D_5ClN_3O$	226.0670	J	J	1.11	1246818-99-4
Benzotriazole D4	$C_6HN_3D_4$	123.0735	J	J	1.30	1185072-03-0
Bromacil-D3	$C_9H_{10}D_3BrN_2O_2$	263.0349	J	J	1.69	2714436-92-5
Simazine-D10	$C_7H_2D_{10}ClN_5$	211.1409	J	N	1.77	220621-39-6
Diuron-D6	$C_9H_4D_6Cl_2N_2O$	238.0547	J	J	1.93	1007536-67-5
DEET_D7	$C_{12}H_{10}D_7NO$	198.1750	J	N	2.50	1219799-37-7
Carbamazepine-D10	$C_{15}H_2D_{10}N_2O$	246.1577	J	N	2.77	132183-78-9
Bezafibrate D6	$C_{19}H_{14}ClNO_4D_6$	367.1457	J	J	2.79	1219802-74-0
Darunavir D9	$C_{27}H_{28}D_9N_3O_7S$	556.2917	J	J	2.82	1133378-37-6
Fipronil 13C4	$C_8{}^{13}C_4H_4Cl_2F_6N_4OS$	439.9521	J	J	4.49	2140327-54-2
Diflufenican-D3	$C_{19}H_8D_3F_5N_2O_2$	397.0929	J	N	5.11	1185009-29-3
Irbesartan D4	$C_{25}H_{24}D_4N_6O$	432.2576	J	J	5.44	1216883-23-6
Telmisartan-D3	$C_{33}D_3H_{27}N_4O_2$	517.2557	J	N	6.04	1189889-44-8

The mix can be obtained as a ready-to-use stock solution as "Pharma-Dx - Mix 22" under the article number " R01243-WaaAN5" from Neochema GmbH (Uwe-Zeidler-Ring 10, 55294 Bodenheim, Germany).

Table B.2: List of isotope-labeled internal standards, eawag (N_{ESI+}= 123, N_{ESI-}= 56)[1]

No.	Name	Sum formula	Retention time [min]
1	2,4-D d3 (-)	$C_8H_3{}^2H_3Cl_2O_3$	9.7
2	2,6-Dichlorobenzamide-3,4,5 d3 (+)	$C_7H_2{}^2H_3Cl_2NO$	5.8
3	5-Methylbenzotriazole d6	$C_7H^2H_6N_3$	6.5
4	Acetyl-sulfamethoxazole d5	$C_{12}H_8{}^2H_5N_3O_4S$	7.0
5	Alachlor d13 (+)	$C_{14}H_7{}^2H_{13}ClNO_2$	12.8
6	Amisulpride d5	$C_{17}H_{22}{}^2H_5N_3O_4S$	5.1
7	Atazanavir d5	$C_{38}H_{47}{}^2H_5N_6O_7$	10.2
8	Atenolol acid d5	$C_{14}H_{16}{}^2H_5NO_4$	4.8
9	Atenolol d7 (+)	$C_{14}H_{15}{}^2H_7N_2O_3$	4.5
10	Atomoxetine d3 (+)	$C_{17}H_{18}{}^2H_3NO$	7.7
11	Atorvastatin d5	$C_{33}H_{30}{}^2H_5FN_2O_5$	11.8
12	Atrazine d5 (+)	$C_8H_9{}^2H_5ClN_5$	9.7
13	Atrazine-2-hydroxy d5	$C_8H_{10}{}^2H_5N_5O$	4.9
14	Atrazine-desisopropyl d5 (+)	$C_5H_3{}^2H_5ClN_5$	5.5
15	Azithromycin d3 (+)	$C_{38}H_{69}{}^2H_3N_2O_{12}$	5.8
16	Azoxystrobin d4 (+)	$C_{22}H_{13}{}^2H_4N_3O_5$	11.8
17	Bentazon d6	$C_{10}H_6{}^2H_6N_2O_3S$	9.4
18	Benzotriazole d4	$C_6H^2H_4N_3$	5.5
19	Bezafibrate d4	$C_{19}H_{16}{}^2H_4ClNO_4$	10.4
20	Bicalutamide d4	$C_{18}H_{10}{}^2H_4F_4N_2O_4S$	11.0
21	Caffeine d9 (+)	$C_8H^2H_9N_4O_2$	5.0
22	Candesartan d5	$C_{24}H_{15}{}^2H_5N_6O_3$	9.3
23	Carbamazepine d8 (+)	$C_{15}H_4{}^2H_8N_2O$	8.4
24	Carbamazepine-10,11-epoxide C13,d2 (+)	$C_{14}{}^{13}CH_{10}{}^2H_2N_2O_2$	7.2
25	Carbendazim d4 (+)	$C_9H_5{}^2H_4N_3O_2$	4.8
26	Cetirizine d8	$C_{21}H_{17}{}^2H_8ClN_2O_3$	8.3
27	Chloridazon d5	$C_{10}H_3{}^2H_5ClN_3O$	6.4
28	Chloridazon-methyl-desphenyl d3	$C_5H_3{}^2H_3ClN_3O$	4.5
29	Chlorotoluron d6 (+)	$C_{10}H_7{}^2H_6ClN_2O$	9.3
30	Chlorpyrifos d10 (+)	$C_9H^2H_{10}Cl_3NO_3PS$	15.9
31	Chlorpyrifos-methyl d6 (+)	$C_7H^2H_6Cl_3NO_3PS$	14.4
32	Citalopram d6 (+)	$C_{20}H_{15}{}^2H_6FN_2O$	7.3
33	Clarithromycin-N-methyl d3 (+)	$C_{38}H_{66}{}^2H_3NO_{13}$	8.4
34	Climbazole d4	$C_{15}H_{13}{}^2H_4ClN_2O_2$	8.4
35	Clofibric acid d4 (-)	$C_{10}H_7{}^2H_4ClO_3$	10.2
36	Clopidogrel carboxylic acid d4 (+)	$C_{15}H_{10}{}^2H_4ClNO_2S$	6.1
37	Clothianidin d3	$C_6H_5{}^2H_3ClN_5O_2S$	6.3
38	Clotrimazole d5 (+)	$C_{22}H_{12}{}^2H_5ClN_2$	8.7
39	Clozapine d8 (+)	$C_{18}H_{11}{}^2H_8ClN_4$	6.5
40	Codeine 13C,d3 (+)	$C_{17}{}^{13}CH_{18}{}^2H_3NO_3$	4.7
41	Cyclophosphamide d4 (+)	$C_7H_{11}{}^2H_4Cl_2N_2O_2P$	7.0
42	Cyprodinil d5 (+)	$C_{14}{}^2H_5H_{10}N_3$	10.7
43	Darunavir d9	$C_{27}H_{28}{}^2H_9N_3O_7S$	10.4
44	Desethylatrazine 15N3 (+)	$C_6H_{10}ClN_2{}^{15}N_3$	6.5
45	Desphenyl chloridazone 15N2 (+)	$C_4H_4ClN{}^{15}N_2O$	2.9
46	Diazepam d5 (+)	$C_{16}H_8{}^2H_5N_2OCl$	10.7
47	Diazinon d10 (+)	$C_{12}H_{11}{}^2H_{10}N_2O_3PS$	14.1
48	Dichlorprop d6 (-)	$C_9H_2{}^2H_6Cl_2O_3$	10.7
49	Diclofenac d4	$C_{14}H_7{}^2H_4Cl_2NO_2$	12.1
50	Diflufenican d3	$C_{19}H_8{}^2H_3F_5N_2O_2$	14.7
51	Dimethenamide d3 (+)	$C_{12}H_{15}{}^2H_3ClNO_2S$	11.7
52	Dimethoate d6 (+)	$C_5H_6{}^2H_6NO_3PS_2$	6.7
53	Diuron d6	$C_9H_4{}^2H_6Cl_2N_2O$	9.8
54	Emtricitabine 13C,15N2 (+)	$C_7{}^{13}CH_{10}FN{}^{15}N_2O_3S$	4.5
55	Epoxiconazole d4 (+)	$C_{17}H_9{}^2H_4ClFN_3O$	11.9
56	Eprosartan d3	$C_{23}H_{21}{}^2H_3N_2O_4S$	6.6
57	Erythromycin 13C2 (+)	$C_{35}{}^{13}C_2H_{67}NO_{13}$	7.4
58	Fenofibrate d6 (+)	$C_{20}H_{15}{}^2H_6ClO_4$	15.9
59	Fipronil 13C2,15N2	$C_{10}{}^{13}C_2H_4Cl_2F_6N_2{}^{15}N_2OS$	13.4
60	Fluconazole d4	$C_{13}H_8{}^2H_4F_2N_6O$	5.9
61	Fluoxetine d5 (+)	$C_{17}H_{13}{}^2H_5F_3NO$	8.4

[1] Eawag - Environmental Chemistry

No.	Name	Sum formula	Retention time [min]
62	Furosemide d5 (-)	$C_{12}H_6{}^2H_5ClN_2O_5S$	8.3
63	Gabapentin D4	$C_9H_{13}{}^2H_4NO_2$	4.7
64	Hydrochlorothiazide 13C,d2	$C_6{}^{13}CH_6{}^2H_2ClN_3O_4S_2$	5.1
65	Ibuprofen d3 (+)	$C_{13}H_{15}{}^2H_3O_2$	12.4
66	Imidacloprid d4	$C_9H_6{}^2H_4ClN_5O_2$	6.5
67	Indomethacin d4	$C_{19}H_{12}{}^2H_4ClNO_4$	12.1
68	Irbesartan d3	$C_{25}H_{25}{}^2H_3N_6O$	8.8
69	Irgarol d9 (+)	$C_{11}H_{10}{}^2H_9N_5S$	9.8
70	Isoproturon d6 (+)	$C_{12}H_{12}{}^2H_6N_2O$	9.7
71	Lamotrigine 13C3,d3 (+)	$C_6{}^{13}C_3H_4{}^2H_3Cl_2N_5$	5.4
72	Levetiracetam d3 (+)	$C_8H_{11}{}^2H_3N_2O_2$	4.8
73	Lidocaine d10 (+)	$C_{14}H_{12}{}^2H_{10}N_2O$	5.3
74	Linuron d6	$C_9H_4{}^2H_6Cl_2N_2O_2$	11.4
75	MCPA d3 (-)	$C_9H_6{}^2H_3ClO_3$	9.8
76	Mecoprop d6 (-)	$C_{10}H_5{}^2H_6ClO_3$	10.6
77	Mefenamic acid d3	$C_{15}H_{12}{}^2H_3NO_2$	13.2
78	Mesotrione d3	$C_{14}H_{10}{}^2H_3NO_7S$	8.8
79	Metalaxyl d6 (+)	$C_{15}H_{15}{}^2H_6NO_4$	9.8
80	Methiocarb d3 (+)	$C_{11}H_{12}{}^2H_3NO_2S$	11.2
81	Methylprednisolone d3 (+)	$C_{22}H_{27}{}^2H_3O_5$	8.4
82	Metolachlor d6 (+)	$C_{15}H_{16}{}^2H_6ClNO_2$	12.8
83	Metolachlor-ESA d11	$C_{15}H_{12}{}^2H_{11}NO_5S$	7.2
84	Metoprolol d7 (+)	$C_{15}H_{18}{}^2H_7NO_3$	5.6
85	Metronidazole d4 (+)	$C_6H_5{}^2H_4N_3O_3$	4.7
86	Metsulfuron-methyl d3	$C_{14}H_{12}{}^2H_3N_5O_6S$	8.8
87	Morphine d3 (+)	$C_{17}H_{16}{}^2H_3NO_3$	4.3
88	N,N-Diethyl-3-methylbenzamide d10 (+)	$C_{12}H_7{}^2H_{10}NO$	9.8
89	N,O-Didesmethyl venlafaxine d3 (+)	$C_{15}H_{20}{}^2H_3NO_2$	5.1
90	N4-acetyl-sulfathiazole d4	$C_{11}H_7{}^2H_4N_3O_3S_2$	5.4
91	Naproxen d3 (+)	$C_{14}H_{11}{}^2H_3O_3$	10.3
92	Nelfinavir d3	$C_{32}H_{42}{}^2H_3N_3O_4S$	8.9
93	Nicosulfuron d6	$C_{15}H_{12}{}^2H_6N_6O_6S$	7.8
94	Octhilinone d17 (+)	$C_{11}H_2{}^2H_{17}NOS$	11.5
95	O-Desmethylvenlafaxine d6 (+)	$C_{16}H_{19}{}^2H_6NO_2$	5.2
96	Oxazepam d5	$C_{15}H_6{}^2H_5ClN_2O_2$	8.8
97	Oxcarbazepine d4 (+)	$C_{15}H_8{}^2H_4N_2O_2$	7.5
98	Paracetamol d4 (+)	$C_8H_5{}^2H_4NO_2$	4.7
99	Phenazone d3 (+)	$C_{11}H_9{}^2H_3N_2O$	5.8
100	Pirimicarb d6 (+)	$C_{11}H_{12}{}^2H_6N_4O_2$	5.9
101	Pravastatin d3 (-)	$C_{23}H_{33}{}^2H_3O_7$	8.1
102	Primidone d5 (+)	$C_{12}H_9{}^2H_5N_2O_2$	5.8
103	Prochloraz d7 (+)	$C_{15}H_9{}^2H_7Cl_3N_3O_2$	11.0
104	Propamocarb free base d7 (+)	$C_9H_{13}{}^2H_7N_2O_2$	4.6
105	Propazine d6 (+)	$C_9H_{10}{}^2H_6ClN_5$	11.0
106	Propiconazole d5 (+)	$C_{15}H_{12}{}^2H_5Cl_2N_3O_2$	13.0
107	Propranolol d7 (+)	$C_{16}H_{14}{}^2H_7NO_2$	6.7
108	Pyrimethanil d5 (+)	$C_{12}H_8{}^2H_5N_3$	9.1
109	Ranitidine d6	$C_{13}H_{16}{}^2H_6N_4O_3S$	4.5
110	Ritalinic acid d10 (+)	$C_{13}H_7{}^2H_{10}NO_2$	5.2
111	Ritonavir d6 (+)	$C_{37}H_{42}{}^2H_6N_6O_5S_2$	12.4
112	Simazine d5 (+)	$C_7H_7{}^2H_5ClN_5$	8.3
113	Sotalol d6	$C_{12}H_{14}{}^2H_6N_2O_3S$	4.5
114	Sulcotrione d3	$C_{14}H_{10}{}^2H_3ClO_5S$	9.0
115	Sulfadiazine d4	$C_{10}H_6{}^2H_4N_4O_2S$	5.1
116	Sulfadimethoxine d4	$C_{12}H_{10}{}^2H_4N_4O_4S$	7.7
117	Sulfamethazine 13C6	$C_6{}^{13}C_6H_{14}N_4O_2S$	5.9
118	Sulfamethoxazole d4	$C_{10}H_7{}^2H_4N_3O_3S$	6.8
119	Sulfapyridine d4	$C_{11}H_7{}^2H_4N_3O_2S$	5.3
120	Sulfathiazole d4	$C_9H_5{}^2H_4N_3O_2S_2$	5.1
121	Tebuconazole d6 (+)	$C_{16}H_{16}{}^2H_6ClN_3O$	12.2
122	Terbuthylazine d5 (+)	$C_9H_{11}{}^2H_5ClN_5$	11.3
123	Terbutryn d5 (+)	$C_{10}H_{14}{}^2H_5N_5S$	9.4
124	Thiamethoxam d3 (+)	$C_8H_7{}^2H_3ClN_5O_3S$	5.7
125	Tramadol d6 (+)	$C_{16}H_{19}{}^2H_6NO_2$	5.6
126	Trimethoprim d9 (+)	$C_{14}H_9{}^2H_9N_4O_3$	4.9
127	Valsartan 13C5,15N	$C_{19}{}^{13}C_5H_{29}N_4{}^{15}NO_3$	10.8

No.	Name	Sum formula	Retention time [min]
128	Valsartan acid d4	$C_{14}H_6{}^2H_4N_4O_2$	7.3
129	Venlafaxine d6 (+)	$C_{17}H_{21}{}^2H_6NO_2$	6.3
130	Verapamil d6 (+)	$C_{27}H_{32}{}^2H_6N_2O_4$	8.1

(+)ESI positive

(-): ESI negative mode

B.2 Standard for retention time standardization and application

Table B.3: List of possible reference standards for RT control and normalization (distribution over the polarity range that can be covered by RP-LC)

Name	Sum formula	logP (log K_{OW})
Metformin	$C_4H_{11}N_5$	-1.36
Chloridazon	$C_{10}H_8ClN_3O$	1.11
Carbetamide	$C_{12}H_{16}N_2O_3$	1.65
Monuron	$C_9H_{11}ClN_2O$	1.93
Metobromuron	$C_9H_{11}BrN_2O_2$	2.24
Chlorobromuron	$C_9H_{10}BrClN_2O_2$	2.85
Metconazole	$C_{17}H_{22}ClN_3O$	3.59
Diazinon	$C_{12}H_{21}N_2O_3PS$	4.19
Quinoxyfen	$C_{15}H_8Cl_2FNO$	4.98
Fenofibrates	$C_{20}H_{21}ClO_4$	5.28

Table B.4: List of substances found in proficiency test B with the number of RTI detections from 6 laboratories with the mean value of the logD deviations and standard deviation

Name	CAS No.	Sum formula	logD (pH 3)	ESI mode	N_{RTI} *)	$\bar{x}$ Δ logD	s Δ logD
Gabapentin	*60142-96-3*	$C_9H_{17}NO_2$	*-2.00*	pos	18	1.4	0.61
				neg	12	1.5	0.73
Metoprolol acid	*56392-14-4*	$C_{14}H_{21}NO_4$	*-1.69*	pos	15	1.1	0.62
				neg	4	1.0	0.01
Propranolol	525-66-6	$C_{16}H_{21}NO_2$	-0.66	pos	15	1.1	0.31
				neg	-	-	-
Hydrochlorothiazide	58-93-5	$C_7H_8ClN_3O_4S_2$	-0.58	pos	10	-0.5	0.18
				neg	14	-0.3	0.27
Caffeine	58-08-2	$C_8H_{10}N_4O_2$	-0.55	pos	17	0.0	0.24
				neg	-	-	-
Clarithromycin	*81103-11-9*	$C_{38}H_{69}NO_{13}$	*-0.26*	pos	16	1.6	0.45
				neg	4	2.1	0.45
Atrazine-2-hydroxy	*2163-68-0*	$C_8H_{15}N_5O$	*0.00*	pos	14	-0.4	0.41
				neg	10	-0.6	0.08
Metamitron	41394-05-2	$C_{10}H_{10}N_4O$	0.24	pos	14	-0.3	0.14
				neg	7	-0.2	0.02
Sulfathiazole	72-14-0	$C_9H_9N_3O_2S_2$	0.93	pos	13	-0.7	0.24
				neg	9	-0.8	0.12
Desethylatrazine	6190-65-4	$C_6H_{10}ClN_5$	1.02	pos	15	-0.8	0.08
				neg	-	-	-
1,2,3-benzo-triazole	*95-14-7*	$C_6H_5N_3$	*1.30*	pos	15	-0.6	0.06
				neg	11	-0.6	0.07
2,4-Dinitrophenol	*51-28-5*	$C_6H_4N_2O_5$	*1.53*	pos	15	-0.2	0.55
				neg	18	-0.1	0.55
4-Methyl-1H-benzotriazole	*29878-31-7*	$C_7H_7N_3$	*1.78*	pos	13	-0.5	0.09
				neg	6	-0.6	0.10
5-Methyl-1H-benzotriazole	*136-85-6*	$C_7H_7N_3$	*1.81*	pos	16	-0.6	0.11
				neg	11	-0.6	0.11
4-Chlorine-benzoic acid	74-11-3	$C_7H_5ClO_2$	2.20	pos	3	-0.5	0.66
				neg	6	-0.3	0.47
N,N-diethyl-toluamide	134-62-3	$C_{12}H_{17}NO$	2.50	pos	15	-0.6	0.86
				neg	-	-	-
Isoproturon	34123-59-6	$C_{12}H_{18}N_2O$	2.57	pos	14	-0.3	0.11
				neg	-	-	-
Mecoprop	7085-19-0	$C_{10}H_{11}ClO_3$	2.85	pos	13	-0.2	0.35
				neg	13	-0.2	0.35
Dimethenamide	87674-68-8	$C_{12}H_{18}ClNO_2S$	2.92	pos	14	-0.1	0.07
				neg	-	-	-
Dinoterb	*1420-07-1*	$C_{10}H_{12}N_2O_5$	*3.09*	pos	12	0.0	0.53
				neg	15	0.5	0.59
Valsartanic acid	164265-78-5	$C_{14}H_{10}N_4O_2$	3.14	pos	18	-1.5	0.44
				neg	18	-1.5	0.44
Metolachlor	51218-45-2	$C_{15}H_{22}ClNO_2$	3.45	pos	16	0.0	0.16
				neg	-	-	-
Bezafibrate	41859-67-0	$C_{19}H_{20}ClNO_4$	3.93	pos	16	-1.4	0.28
				neg	16	-1.4	0.28
Gemfibrozil	*25812-30-0*	$C_{15}H_{22}O_3$	*4.37*	pos	4	0.1	0.57
				neg	5	0.2	0.52

*) from 6 laboratories

Appendix C. Methodical

C.1 Examples of LC methods

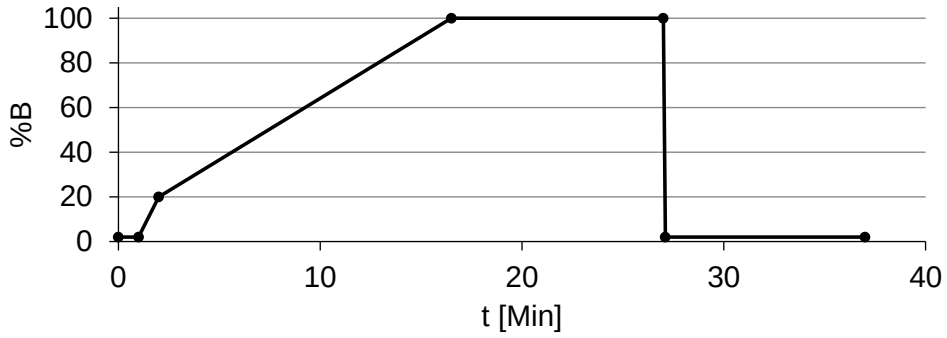
Two example methods for possible chromatographic separation methods are given below.

Method A:

Eluents			
	A: MilliQ + 0.1% v/v formic acid		
	B: Acetonitrile + 0.1% v/v formic acid		
Injection volume	95 µL sample + 5 µL isotope-labeled standard mix		
Column temperature	40°C		
Flow rate	0.3 mL/min		
Pillar	Agilent Zorbax Eclipse Plus C18		
	Narrow Bore RR	2.1x150 mm	3.5 µm
	PN:	959763-902	
Pre-column	Phenomenex Cartridge Holder		
	C18 4x2.0 mm ID		
	PN:	AJO-4286	

Gradient

%B	2	2	20	100	100	2	2
t [min]	0	1	2	16,5	27	27,1	37

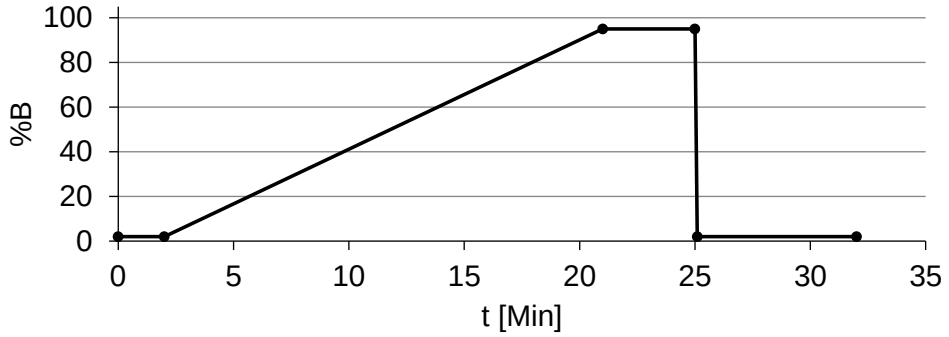


Method B:

Eluents	A: MilliQ + 0.1% v/v formic acid B: Acetonitrile + 0.1% v/v formic acid
Injection volume	95 µL sample + 5 µL isotope-labeled standard mix
Column temperature	40°C
Flow rate	0.6 mL/min
Pillar	Restek Ultra Aqueous C18 250 x 4.6 mm 5 µm Cat: 9178575
Pre-column	Restek Ultra AQ C18 10 x 4 mm Cat: 917850210

Gradient

%B	2	2	95	95	2	2
t [min]	0	2	21	25	25,1	32



C.2 Examples of MS methods

Two examples of possible mass spectrometric detection methods using time-of-flight mass spectrometers or Orbitrap systems are given below.

Table C.1: Example of an MS method (TOF system)

Source parameters	
Gas Flows	Gas 1: 35 psi Gas 2: 45 psi Curtain gas: 40 psi Collision gas: 6/medium
Temperature	550 °C
ISVF	5500 V (+) -4500 V (-)
Declustering potential	60 V (+) -100 (-)
TOF-MS scan parameters	
Mass Range	MS: 100 - 1200 Da TOF-MS: 250 ms
MS² parameters	
Mass Range	30 - 1200 Da
Collision Energy	40 eV (+) -40 eV (-)
Collision Energy Spread	20 eV
MS² Acquisition in IDA or SWATH mode	
IDA Triggering	
Accumulation Time	65 ms
Max number of MS² per cycle	12
Minimum intensity	100 cps
Exclude isotopes	Within 4 Da
Mass Tolerance	5 ppm
Include/Exclude List	None
Dynamic Background subtract	On
SWATH	
Accumulation Time	50 ms
Mass range	100 - 1200 Da
Number of SWATH windows	16

Table C.2: Example of an MS method (Orbitrap system)

Source parameters	
Gas Flows	Sheath Gas: 40
	Aux gas flow: 15
	Sweep gas: 50
Temperature	Capillary:350 °C
	Aux gas: 400 °C
Spray Voltage	3500 V
MS Scan parameters	
Mass Range	Full MS: 120 - 1200 *m/z*
Resolution	30,000
Microscans	1
Maximum inject time	50 ms
Full MS / dd-MS² (Top N)	
Full MS	
Resolution	120,000
AGC Target	$3e^6$
Maximum IT	100 ms
Scan Range	120 - 1200 *m/z*
dd-MS²	
Resolution	15,000
AGC Target	$1e^5$
Maximum IT	50 ms
Loop count	5
Isolation window	1.3 *m/z*
Fixed first mass	50.0 *m/z*
(N)CE / stepped N(CE)	Nce: 80
dd Settings	
Minimum AGC target	$8.00e^3$
Apex trigger	3 to 10 s
Charge Exclusion	-
Peptide Match	Preferred
Exclude isotopes	On
Dynamic exclusion	15.0 s

C.3 Blank value measurements

For the two example methods, the total ion current chromatograms for electrospray ionization are both positive and negative below. The intensity axis is scaled the same for all chromatograms.

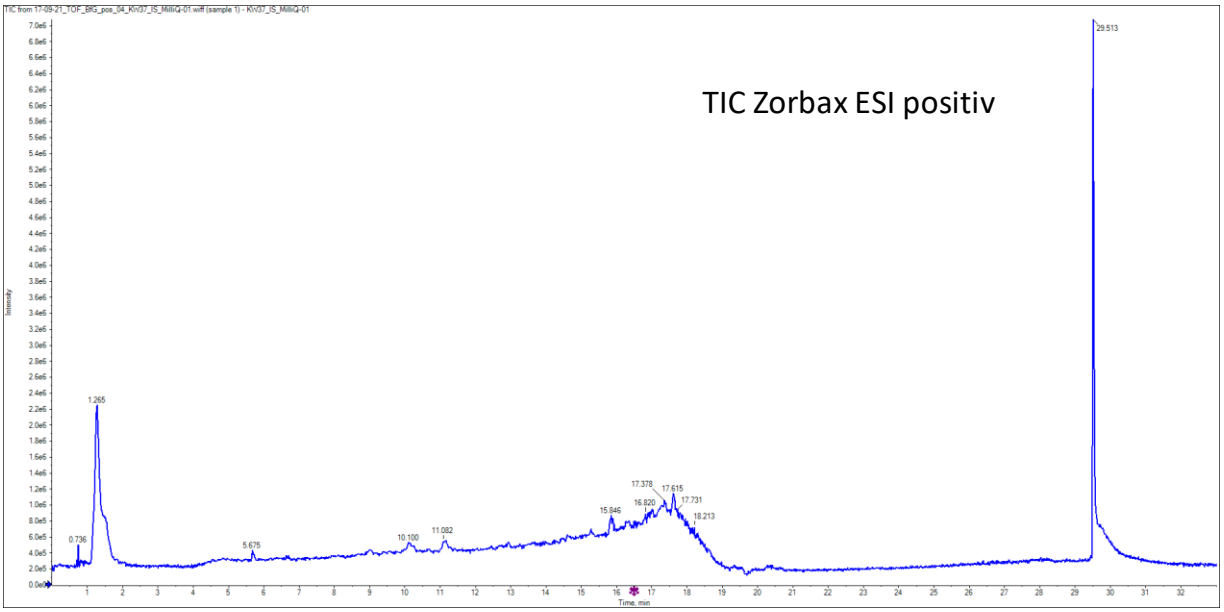


Figure C-1: Total ion current chromatogram LC method A (C.1); electrospray positive

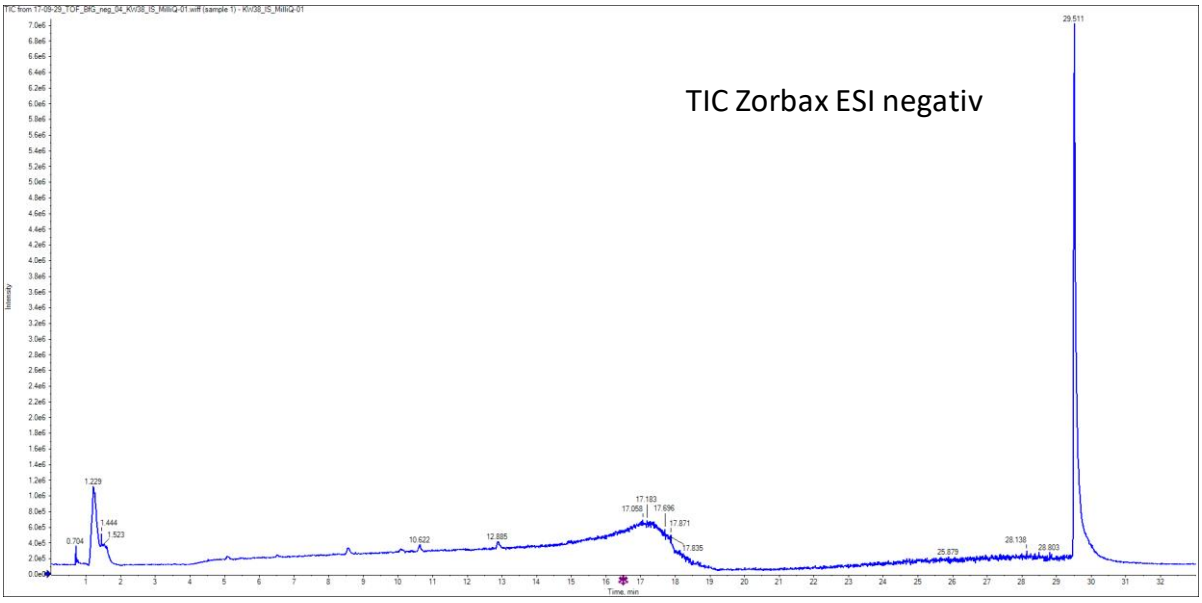


Figure C-2: Total ion current chromatogram LC method A (C.1); electrospray negative

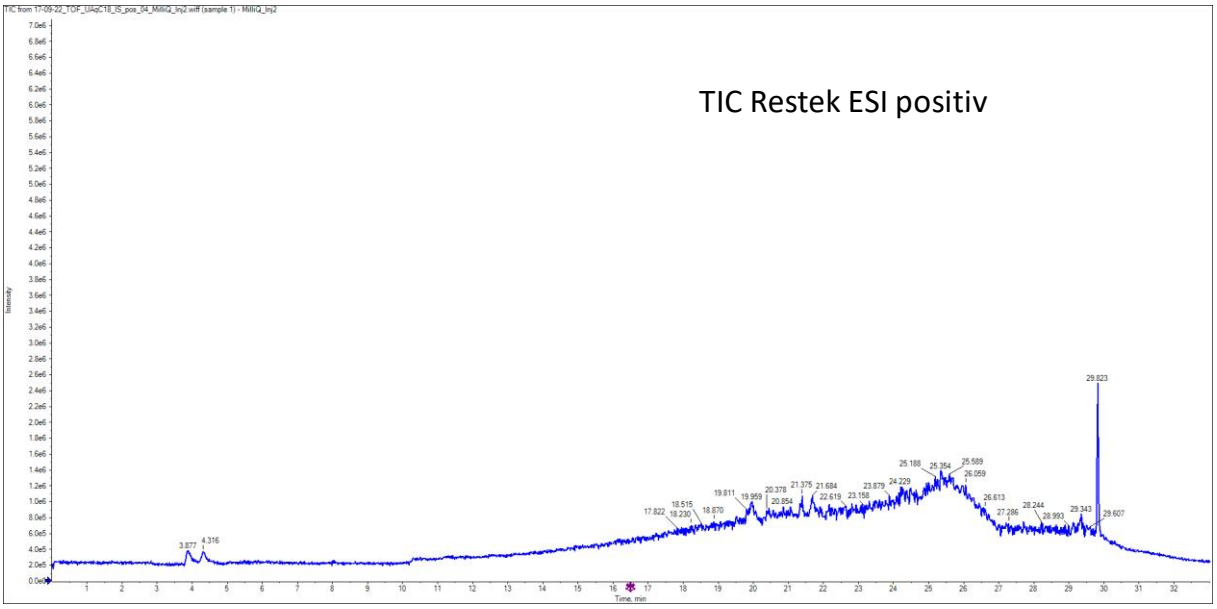


Figure C-3: Total ion current chromatogram LC method B (C.1); electrospray positive

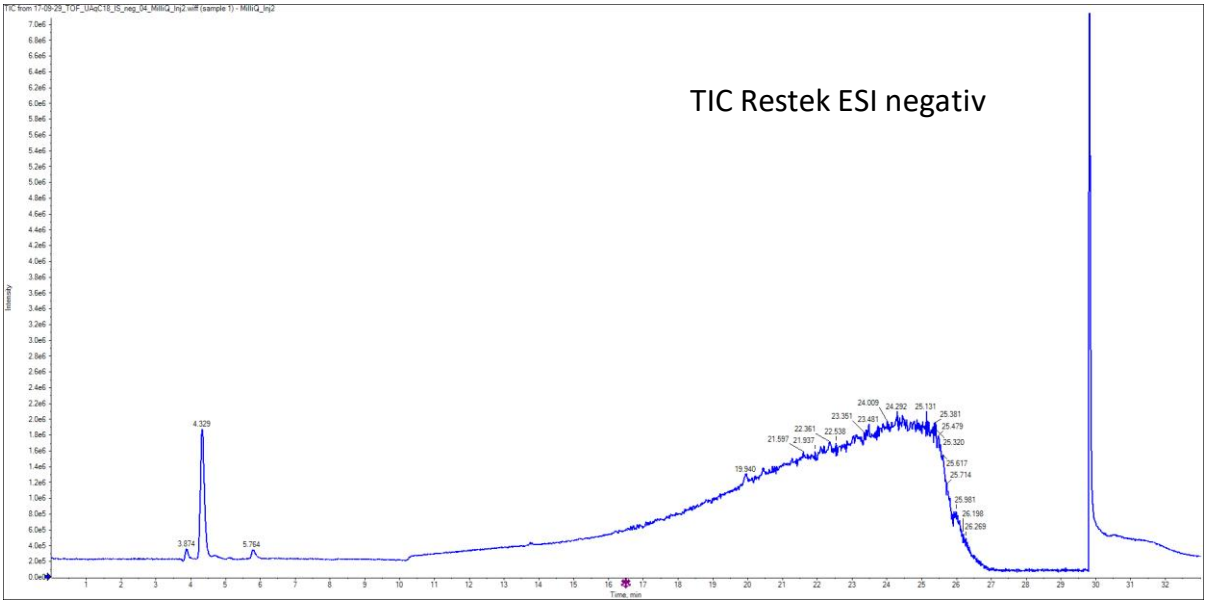


Figure C-4: Total ion current chromatogram LC method B (C.1); electrospray negative

C.4 Retention time-mass plot of blank values

The features detected in the respective blank value are compared as a retention time-mass plot for ESI+ and ESI-. The red dots are the isotope-labeled internal standards used. The internal standards should be evenly distributed over the measured mass and polarity range.

The retention time-mass plots shown in Figure C-5 clearly show different pictures for methods A and B, mainly due to the different stationary phases of the separation column.

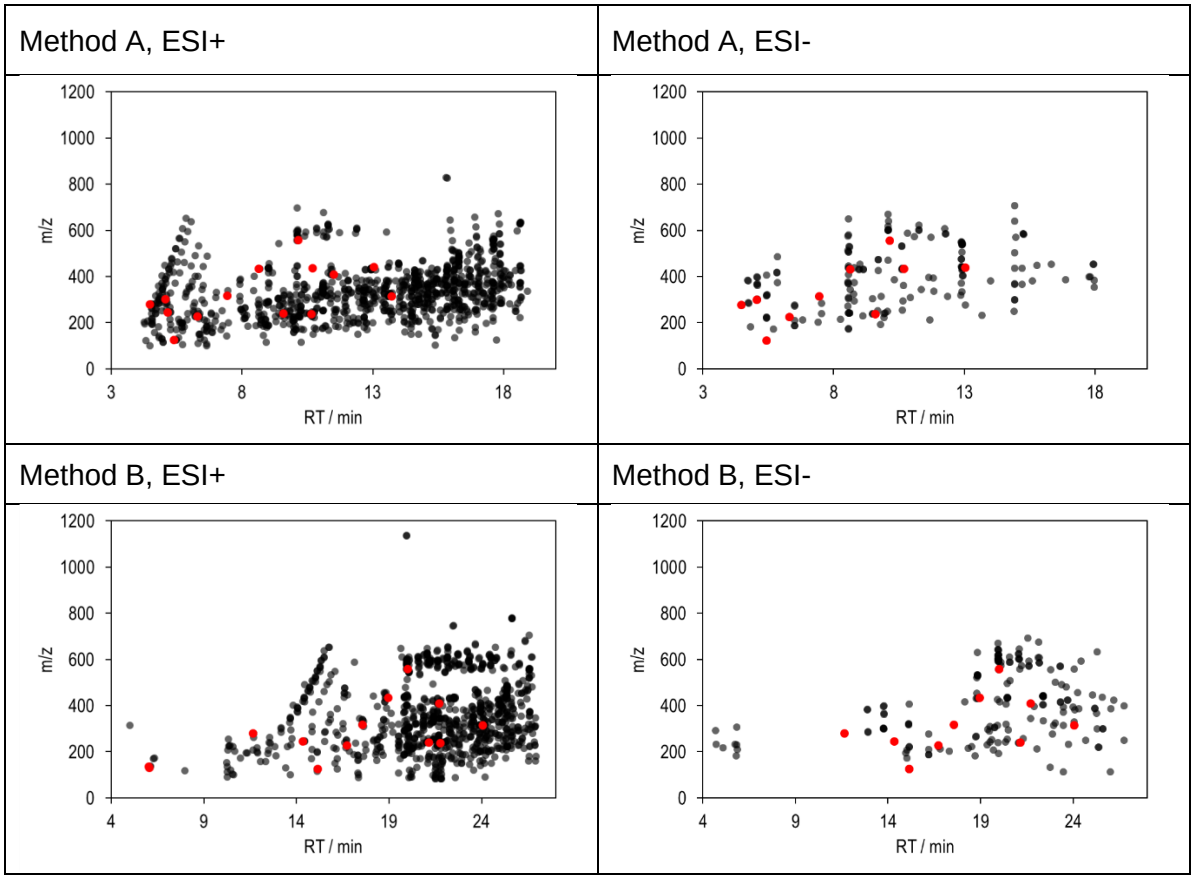


Figure C-5: *Mass RT scatterplots ("point clouds") of the two example methods A and B, each in ESI positive and negative mode*

Appendix D. Measurement technology

D.1 HRMS mass spectrometer

The latest development in ion trap mass spectrometers is the Orbitrap. The ion trap contains a central, spindle-shaped electrode. The ions are introduced into the Orbitrap radially to this electrode and move on circular paths (orbits) around the central electrode due to electrostatic attraction. As the ions are not introduced in the middle of the chamber but in a decentralized manner, they oscillate simultaneously along the axis of the central electrode. The frequency of this oscillation generates signals in detector plates, which are converted into the corresponding *m/z* ratios by Fourier transformation.

A time-of-flight mass spectrometer (TOF-MS) consists of a tube under a vacuum with a high-speed detector at the end. In principle, TOF devices use the fact that the ions all have the same energy when they enter the analyzer and that lighter ions are faster than heavier ions when an accelerating voltage is applied. Therefore, light ions reach the detector earlier than heavy ions when flying through a field-free space (flight tube). In practice, devices with ion mirrors or reflections, in which the flight path is increased by an additional electric field at the end of the original flight direction, have proven their worth. This technique also enables further focusing, which minimizes the variance in the speed of the ions due to the Doppler effect. The length of the flight path is decisive for the resolution of the mass spectrometer.

Orbitrap

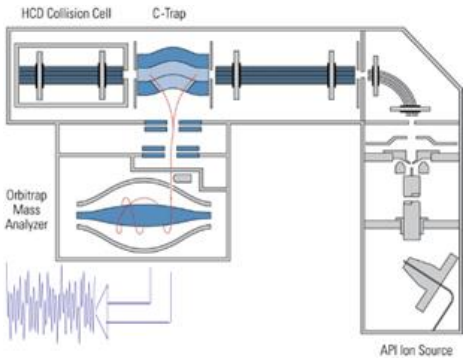


Image source:
Thermo Fischer Scientific

Time-of-flight mass spectrometer (TOF)

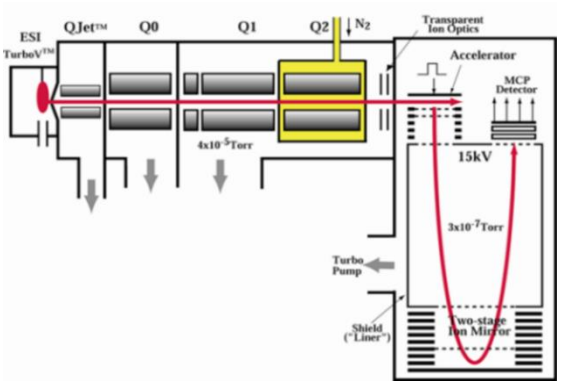


Image source: Sciex®

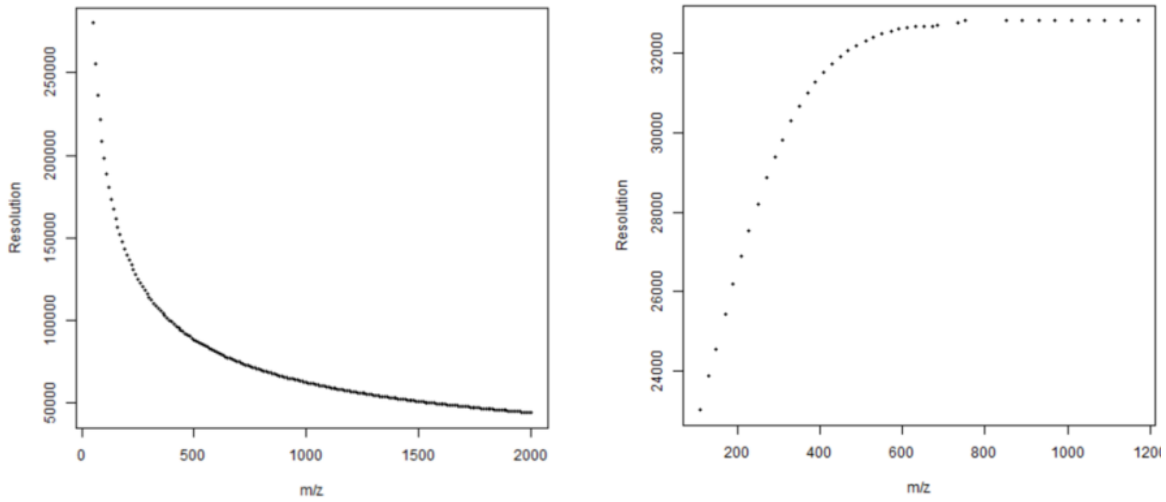


Figure D-1: *Schematic layout of the Orbitrap mass spectrometer (left) and time-of-flight mass spectrometer (right) with their resolution as a function of mass range (bottom)*

Appendix E. System stability

E.1 Chromatography

Reproducibility of the retention time

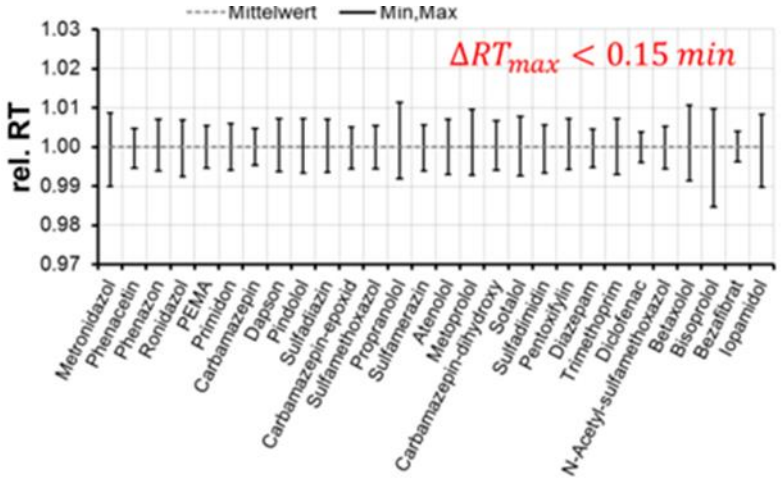


Figure E-1: Retention time stability over N= 134 (over a period of 10 months)

E.2 Mass spectrometry

Long-term stability of sensitivity

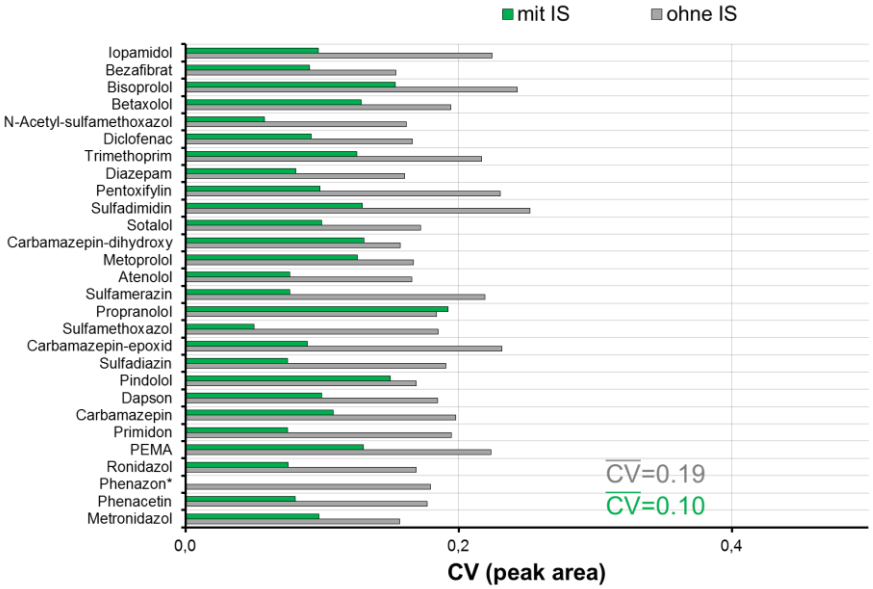


*Figure E-2: Stability of device sensitivity over a period of 10 months (N = 134) without (grey) and with (green) internal standardization (*phenazone as IS)*

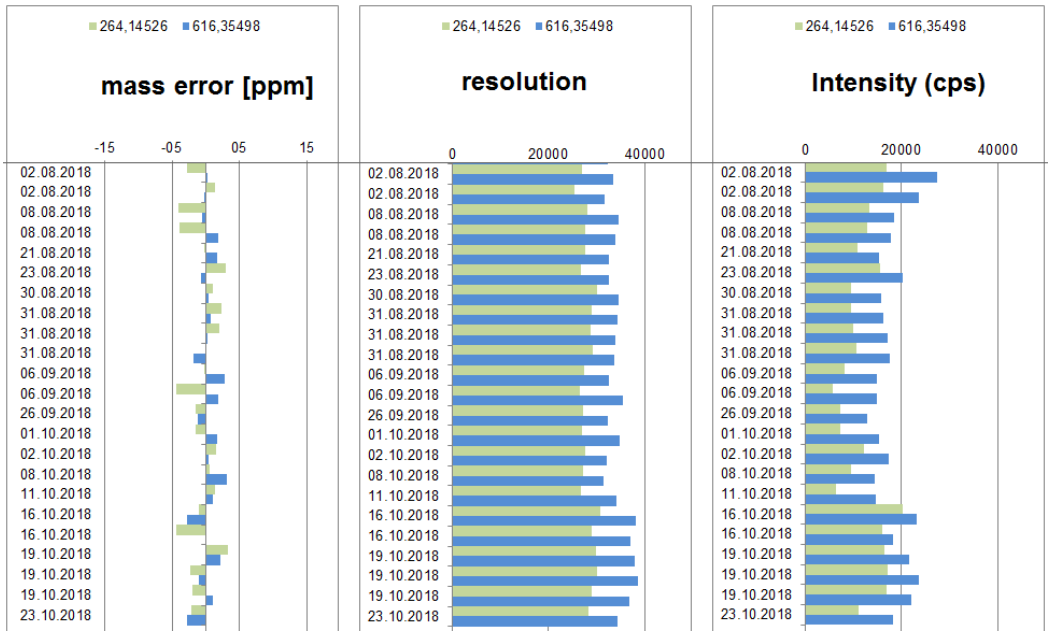


Figure E-3: *Documentation option for monitoring MS performance via mass accuracy, resolution and sensitivity*

Appendix F. Data evaluation

F.1 Adjustment of intensity-dependent parameters for peak extraction using the example of the "Noise Threshold" of the MarkerView™-software (SCIEX)

Multiple determinations of an aliquot of a wastewater treatment plant effluent sample (QA sample) spiked with 64 substances at different times within one year resulted in different sensitivity levels of the measuring device (LC-HRMS). The previously optimized value for the "Noise Threshold" of 100 (positive ionization) or 75 (negative ionization), therefore, did not produce satisfactory results for peak finding (Figure F-2). In addition to higher measurement signals for the real features, an improvement in sensitivity also increases noise. The average noise (median) of all doped substances from the control sample was determined for each measurement to adjust the noise threshold. A noise threshold was calculated from each of these values based on the values of the optimization measurements. The plot of "Noise" against "Noise Threshold" resulted in a linear relationship, the equation of which can be used for further adjustments (Figure F-1).

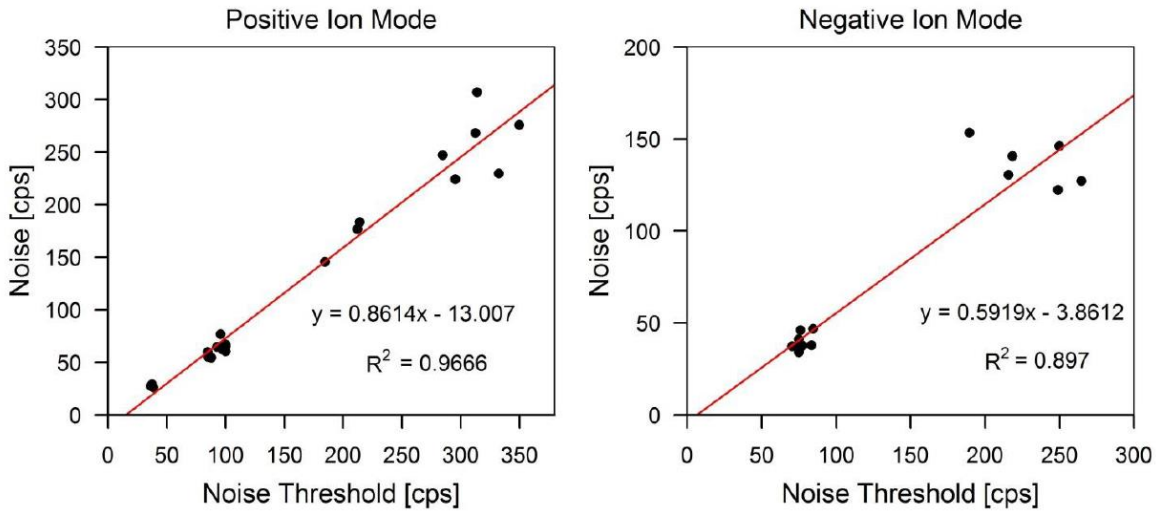


Figure F-1: Correlation between "Noise" and the calculated "Noise Threshold"

The use of these adjusted values for the "Noise Threshold" showed that the proportion of false positives (FPs) of the features corresponded to that of the original optimization (Figure F-2). The adjustment via the noise median thus works very well, but the total number of features varied if a different value had to be used for the "Noise Threshold." With a higher sensitivity, additional features with low intensity can also be detected, which cannot be detected with a lower sensitivity. With an evaluation approach that is only based on the number of features, the comparability of results is only given if the sensitivity differences between different measurement series are not too high.

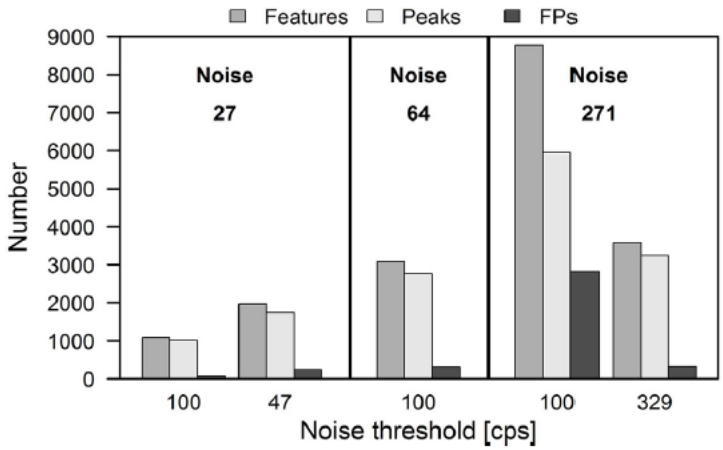


Figure F-2: *Change in the number of features, true peaks, and false positives (FPs) due to the noise threshold used (100 cps and calculated value from the fitting equation) for the measurements (positive ionization) of a spiked WWTP effluent sample for three different sensitivity levels of the instrument. Left: LC-HRMS with low sensitivity, middle: LC-HRMS during optimization, right: LC-HRMS with higher sensitivity. For further details see [2].*

Appendix G. Adduct formation when using an ESI source

G.1 Adducts and in-source *fragments*

Table G.1: Examples of detected adducts and in-source fragments of known substances

Type	split / added elements	Polarity	Description	Mass difference compared to $[M+H]^+$ or $[M-H]^-$	Example compounds
Adduct	+O	both	Addition of oxygen	15.99491	2-mercaptobenzoxazole, 2-mercaptobenzothiazole
Adduct	$+NH_4$	positive	Addition of ammonium	17.02654	Diatrizoate. ethofumesate. iopromide
Adduct	+Na	both	Addition of sodium	21.98194	pos: carbamazepine. metolachlor / neg: Valsartan. Olmesartan
Adduct	+HCl	negative	Addition of HCl	35.97667	Ethidimuron. dimefuron. methoxyfenozide
Adduct	+K	positive	Addition of potassium	37.95588	Azoxystrobin. dimoxystrobin. praziquantel
Adduct	$+C_2H_8N$	positive	Addition of ethylamine	45.05784	Dimethoate. tetraglym. dimefuron. metalaxyl
Adduct	$+CH_2O_2$	negative	Addition of formic acid	46.00548	Flecainide. aliskiren. fluconazole
Adduct	$+C_2H_4O_2$	negative	Addition of acetic acid/ Sodium cluster	60.02113	-
Adduct	$+HNO_3$	negative	Addition of nitrate	62.99564	Clothianidin. Fluconazole
Adduct	$+NaCH_2O_2$	negative	Addition of formic acid/ Sodium cluster	67.98743	Penoxsulam. diphe-nylphosphinicacid. haloxyfop.
Adduct	$+NaC_2H_4O_2$	negative	Addition of acetic acid/ Sodium cluster	83.0109	-
Adduct	$+NaNO_3$	negative	Addition of nitrate/ Sodium cluster	84.97814	Bromacil. Chlorthanonil R611965
Fragment	$-C_7H_8N_2O_4S$	positive		-216.0210	Metazachlor metabolite BH 479 9
Fragment	$-C_{10}H_{14}O_4$	positive		-198.0905	Kresoxim-methyl
Fragment	$-C_5H_6O_4N_2S$	positive		-190.0048	Metazachlor metabolite 479M008
Fragment	$-C_9H_{11}O_4$	positive		-183.06554	Kresoxim-methyl
Fragment	$-C_6H_8O_2N_2S$	positive		-172.0312	Metazachlor metabolite BH 479 11
Fragment	$-C_8H_8O_3$	positive		-152.04789	Metabolites 505M08 and 505M09
Fragment	$-C_6H_8O_3$	positive		-152.0472	Kresoxim-methyl
Fragment	$-C_5H_4O_3N_2$	positive		-140.02274	Metazachlor metabolite NOA409045
Fragment	$-C_4H_8O_5$	positive		-136.03772	Metalaxyl metabolite CGA 108906
Fragment	$-C_2O_2F_9$	negative		-127.00069	ADONA

Type	split / added elements	Polarity	Description	Mass difference compared to $[M+H]^+$ or $[M-H]^-$	Example compounds
Fragment	$-C_7H_5ON$	negative		-119.03711	Carbetamide
Fragment	$-C_3H_2O_5$	positive		-117.99077	Metolachlor metabolite CGA 357704
Fragment	$-C_7H_8O$	positive		-108.05737	Kresoxim-methyl
Fragment	$-C_3H_9O_3N$	positive		-107.05879	Metabolites 505M08 and 505M09
Fragment	$-C_2H_2O_3S$	negative		-105.97301	Dimethenamide metabolite M31. Metazachlor metabolite CGA 368208
Fragment	$-C_3H_4O_4$	negative		-104.01151	Dimethenamide metabolite M23
Fragment	$-C_3H_8O_3$	positive		-92.04721	Kresoxim-methyl
Fragment	$-C_2H_6O_3$	negative		-90.03224	Metalaxyl metabolite CGA 108906
Fragment	$-C_5H_{11}ON$	positive		-89.08406	Diphenhydramine
Fragment	$-C_5H_{12}O$	positive		-88.08882	Pendimethalin
Fragment	$-C_3H_5O_2$	positive		-88.05298	Metolachlor metabolite CGA 50267
Fragment	$-C_2O_4$	negative		-87.98021	Quinmerac metabolite BH 518-2
Fragment	$-C_2H_2O_2N_2$	negative		-86.01218	Thiacloprid metabolite M30
Fragment	$-C_2H_3ON_3$	negative		-85.02816	Tritosulfuron metabolite M635H003
Fragment	$-SO_3$	positive	Separation of SO_3	-79.95682	Sitagliptin-N-sulfate
Fragment	$-C_2H_4O_3$	positive		-76.01596	Kresoxim-methyl. metolachlor Metabolite CGA 37735
Fragment	$-C_3H_5O_2$	positive		-73.0295	Metolachlor metabolite CGA 50267
Fragment	$-C_3H_4O_2$	negative		-72.02058	Mecoprop. Fenoprop. Fluziprop
Fragment	$-C_2O_3$	negative		-71.98419	Dimethenamide metabolite M23
Fragment	$-C_5H_{10}$	positive		-70.07825	Pendimethalin
Fragment	$-C_3H_4N_2$	positive		-68.03745	Prochloraz. Metazachlor metabolite 479M004. Metazachlor metabolite 479M008
Fragment	$-C_5H_6$	positive		-66.04641	Propyzamide
Fragment	$-CH_4O_3$	positive		-64.01605	2-OH-ibuprofen
Fragment	$-C_2H_4O_2$	positive		-60.02168	Metalaxyl metabolite CGA 108906
Fragment	$-C_2H_2O_2$	both		-58.00493	Kresoxim-methyl. Metolachlorometabolite CGA 37735
Fragment	$-C_2H_3ON$	both		-57.02146	DCPMU. Carbofuran. Carbaryl
Fragment	$-C_4H_8$	positive		-56.0626	Bromacil. terbuthylazine. bupropion. methoxyfenozide
Fragment	$-C_3H_4O$	negative		-56.0256	Ketoprofen
Fragment	$-3 \cdot H_2O$	positive	3-fold water splitting	-54.03168	Prednisolone
Fragment	$-CH_6O_2$	positive		-50.03733	Dimethachlor metabolite SYN 530561
Fragment	$-CH_5ON$	positive		-47.03711	Kresoxim-methyl

Type	split / added elements	Polarity	Description	Mass difference compared to $[M+H]^+$ or $[M-H]^-$	Example compounds
Fragment	$-C_2H_6O$	positive		-46.04241	Mefenpyr-diethyl. fenoxycarb. ethofumesate. pethoxamide
Fragment	$-CH_4ON$	positive		-46.02929	Levetiracetam
Fragment	$-CH_2O_2$	both		-46.00548	Naproxen. Ibuprofen
Fragment	$-CO_2$	negative		-43.98986	Diatrizoate. N-methyl-pregabalin
Fragment	-CHON	negative		-43.00581	DCPU. Metabolite M635H001
Fragment	$-C_3H_6$	positive		-42.0475	Flufenacet metabolite AZ14777
Fragment	$-2 \cdot H_2O$	positive	2-fold water splitting	-36.02112	Prednisolone
Fragment	-Cl	positive	Separation of chlorine	-34.9683	3.4-Dichloroaniline
Fragment	$-CH_4O$	both		-32.02622	Dimethenamid. metolachlor. oxfend-azole
Fragment	$-CH_5N$	positive		-31.04219	Sertraline
Fragment	$-CH_2O$	positive		-30.01111	Topramezone metabolite M670H05
Fragment	-HF	negative	Separation of fluorine	-20.00623	Diflubenzuron
Fragment	$-H_2O$	both	Water separa-tion	-18.01056	pos: 10.11-dihydroxy-10.11-dihydro-carbamazepine. gabapentin / neg: diclofenac. PFBA. diatrizoate
Fragment	$-NH_4$	positive		-17.02654	Levetiracetam. amoxicillin
Fragment	$-CH_4$	positive		-16.0313	1.2-Dihydro-2.2.4-trimethyl quinoline
Fragment	-O	positive	Separation of oxygen	-15.99491	Ranitidine-N-oxide. 5-chloro-2-mercaptobenzoxazole

Further adducts. in-source *fragments.* or typical blank values and impurities in the LC-(HR)MS are described in the literature [55].

Appendix H. Statistical methods

The following Table H.1 is a selection of methods used to statistically evaluate NTS data. It should be noted that each method has specific strengths and weaknesses, and it is therefore always necessary to check which method could be used for the specific analytical question. In addition to this selection of methods, there are many other known methods, which is why it is also recommended to carry out a comprehensive literature search if necessary.

Table H.1: Overview of exemplary statistical methods for analyzing NTS data.

Method	**Study (reference)**	**Basic principle**	**Applicable question**
PCA (Principal Component Analysis)	[1,2]	Reduction of dimensionality with maximization of variance	Exploration of the data structure Grouping of similar samples and features Identification of extremes Feature prioritization
HCA (Hierarchical Cluster Analysis)	[3]	Restructuring of samples using a feature-based similarity metric	Exploration of the data structure Grouping of similar samples Identification of extremes
PLS-DA (Partial Least Squares Discriminant Analysis)	[4]	Classification and maximization of variance between groups	Identification of features that best differentiate groups
OPLS-DA (Orthogonal Partial Least Squares Discriminate Analysis)	[5]	Classification and maximization of variance between groups with separate consideration of differences between and within groups	Identification of features that best differentiate groups
ML-PLS-DA (Multi Level Partial Least Squares Discriminant Analysis)		Classification and maximization of variance between subgroups, taking into account higher-level groups	Identification of features that best differentiate (sub)groups. e.g., regional or seasonal comparisons

[1] Purschke, K., Zoell, C., Leonhardt, J., Weber, M., & Schmidt, T. C. (2020). Identification of unknowns in industrial wastewater using offline 2D chromatography and non-target screening. Science of the Total Environment, 706, 135835.

[2] López-Doval, J. C., Montagner, C. C., de Alburquerque, A. F., Moschini-Carlos, V., Umbuzeiro, G., & Pompêo, M. (2017). Nutrients, emerging pollutants and pesticides in a tropical urban reservoir: Spatial distributions and risk assessment. Science of the Total Environment, 575, 1307-1324.

[3] Schollée, J. E., Bourgin, M., von Gunten, U., McArdell, C. S., & Hollender, J. (2018). Non-target screening to trace ozonation transformation products in a wastewater treatment train including different post-treatments. Water Research, 142, 267-278.

[4] Samanipour, S., Kaserzon, S., Vijayasarathy, S., Jiang, H., Choi, P., Reid, M. J., & Thomas, K. V. (2019). Machine learning combined with non-targeted LC-HRMS analysis for a risk warning system of chemical hazards in drinking water: A proof of concept. Talanta, 195, 426-432.

[5] Vanryckeghem, F., Huysman, S., Van Langenhove, H., Vanhaecke, L., & Demeestere, K. (2019). Multi-residue quantification and screening of emerging organic micropollutants in the Belgian Part of the North Sea by use of Speedisk extraction and Q-Orbitrap HRMS. Marine Pollution Bulletin, 142, 350-360.

Method	Study (reference)	Basic principle	Applicable question
MOCA (Multi-block Orthogonal Component Analysis)		Classification and maximization of variance between groups across multiple data sets with separate consideration of differences between and within data sets	Identification of features that best differentiate groups
ASCA (ANOVA-Simultaneous Component Analysis)	[1]	Analysis of the influence of several factors on the data set	Evaluation of the influence of experimental factors
PARAFAC2 (Parallel Factor Analysis 2)	[2]	Pattern recognition in three-way data (e.g., time, region, feature intensity), where one dimension may vary	Time series analysis concerning feature intensity profiles. e.g., to identify trends for forecasting or early warning systems
FBMN (Feature-Based Molecular Networking)	[3]	Network creation based on MS and MS^2 spectrum similarity	Identification of unknown substances
MCR-ALS (Multivariate Curve Resolution - Alternating Least Squares)	[4,5]	Bi-linear regression based on *m/z* related elution profiles	Componentization alternative to the creation of componentized features that does not require feature detection

[1] Hohrenk-Danzouma, L. L., Vosough, M., Merkus, V. I., Drees, F., & Schmidt, T. C. (2022). Non-target analysis and chemometric evaluation of a passive sampler monitoring of small streams. Environmental Science & Technology, 56(9), 5466-5477.

[2] Nielsen, N. J., Christensen, P., Poulsen, K. G., & Christensen, J. H. (2023). Investigation of micropollutants in household waste fractions processed by anaerobic digestion: target analysis, suspect-and non-target screening. Environmental Science and Pollution Research, 30(16), 48491-48507.

[3] Oberleitner, D., Schmid, R., Schulz, W., Bergmann, A., & Achten, C. (2021). Feature-based molecular networking for identification of organic micropollutants including metabolites by non-target analysis applied to riverbank filtration. Analytical and Bioanalytical Chemistry, 413, 5291-5300.

[4] Hohrenk, L. L., Vosough, M., & Schmidt, T. C. (2019). Implementation of chemometric tools to improve data mining and prioritization in LC-HRMS for nontarget screening of organic micropollutants in complex water matrices. Analytical Chemistry, 91(14), 9213-9220.

[5] Khatoonabadi, R. L., Vosough, M., Hohrenk, L. L., & Schmidt, T. C. (2021). Employing complementary multivariate methods for a designed nontarget LC-HRMS screening of a wastewater-influenced river. Microchemical Journal, 160, 105641.

Appendix I. Workflow

I.1 Example of a typical screening workflow

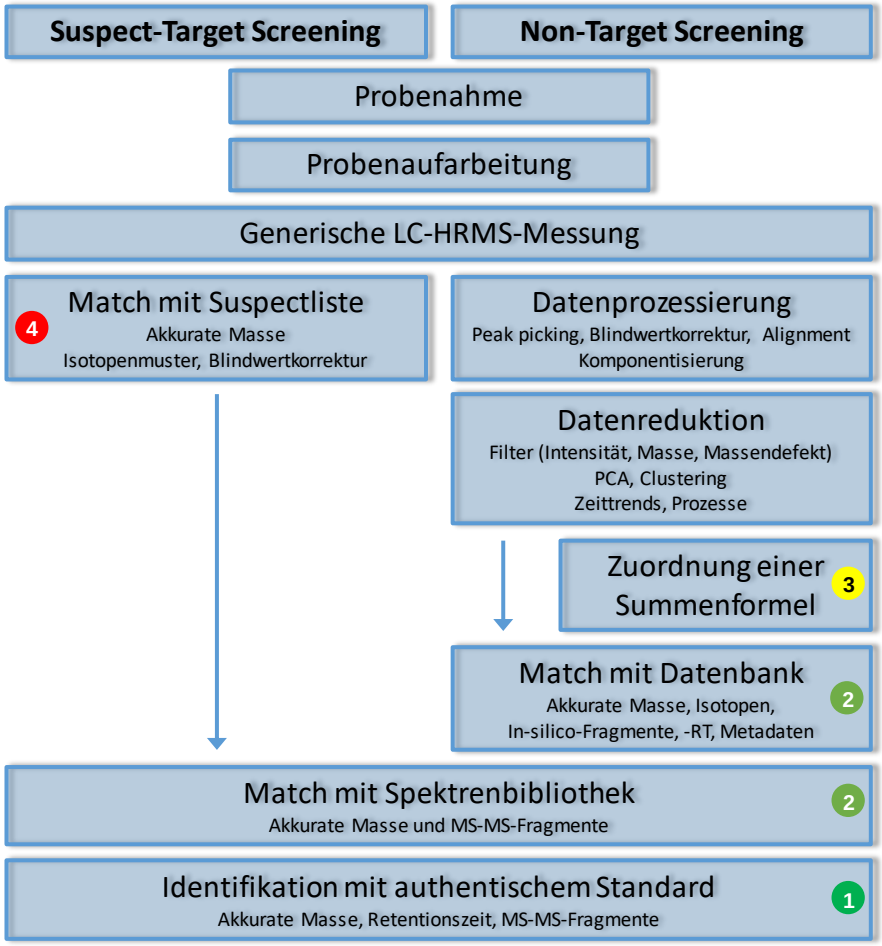


Figure I-1: Example workflow for suspect and non-target screening. including the categorization for the identification of substances (see also 10.2.1)

Further examples of workflows can be found in the literature [1].

Appendix J. Validation

J.1 Sample comparison

For example. feature comparison of repeat measurements can be carried out to validate the precision, accuracy, and measurement uncertainty parameters. After peak finding and alignment, the determined peak intensities are compared (fold-change f_c). Figure J-1 shows the possible intensity ratios in principle. In the "green range", the features were detected in both samples, and their intensity ratios fluctuate within a defined range, for example. $0.5 \leq f_c \leq 2.0$. In the "red range", the more extreme ratios range up to not detected in one or the other measurement. The highest intensity of the peak comparison is plotted in a diagram against the fold-change (see Figure J-2 and Figure J-3).

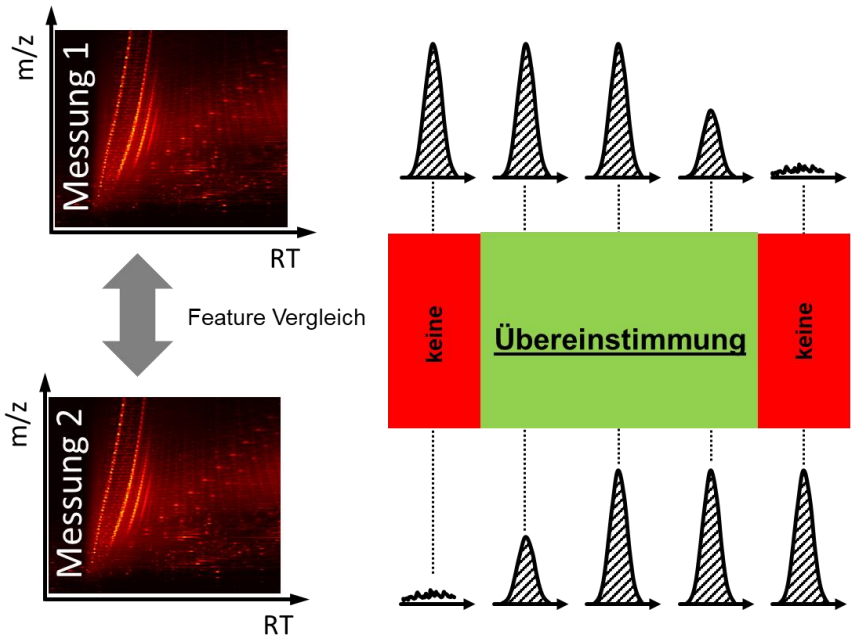


Figure J-1: Evaluation of the sample comparison based on the ratio of the peak intensities

Figure J-2 and Figure J-3 show two examples of the results of a sample comparison. The "green range" can be defined in a case-specific method based on experience with the measuring system. Such a comparison between the measurements of different samples can be made using the internal standards added. This makes it possible to recognize matrix influences during the measurement and consider them in the evaluation.

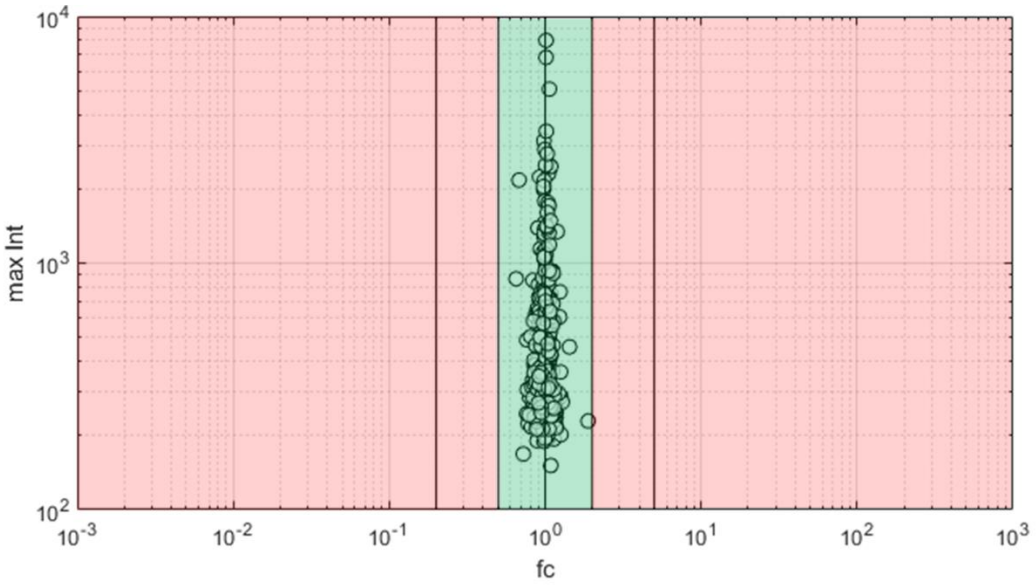


Figure J-2: Exemplary result of a repeat measurement with excellent reproducibility

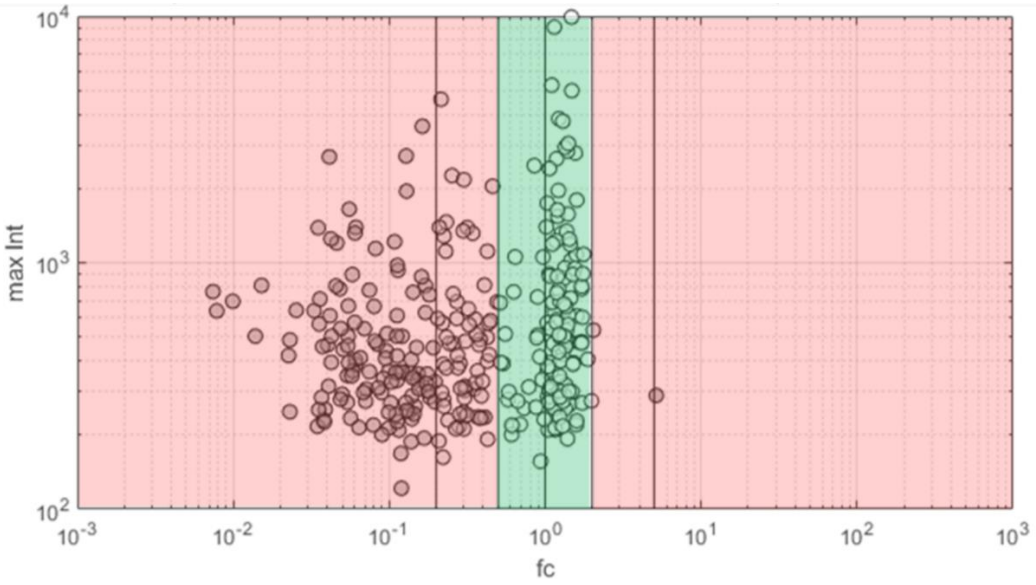


Figure J-3: Exemplary result of a repeat measurement with a very poor reproducibility

J.2 Classification

Non-target screening is an essential tool for classifying samples. This makes it possible to assign samples to different classes based on the pattern of signals, even without identifying individual substances. The confusion matrix Figure J-4 can be used to validate the classification method.

		Realität		
		positiv (A+)	negativ (A-)	Summe
Analyse	positiv (A+)	richtig-positiv (TP)	falsch-positiv (FP)	PP = TP + FP
	negativ (A-)	falsch-negativ (FN)	richtig-negativ (TN)	PN = FN + TN
	Summe	RP = TP + FN	RN = FP + TN	N = PP +PN

Figure J-4: Confusion matrix (truth matrix
PP: predicted positives; PN: predicted negatives; N: total number of samples.
RP: real positives; RN: real negatives

The following key figures can be calculated from the confusion matrix:
(It should be noted that these terms have a different interpretation here than for validation in target analytics).

Sensitivity: This is also referred to as the true-positive rate (recall) and indicates the proportion of positive samples that are actually positive.

$$TPR\ (Recall. Sensitivity) = \frac{TP}{TP + FN}$$

Sensitivity is a key factor in the detection of positive cases.

Specificity: This is also known as the true-negative rate and indicates the proportion of negative samples that are actually negative.

$$TNR(Specificity. Selectivity) = \frac{TN}{TN + FP}$$

Precision: This is the relative frequency of correctly classified positive samples.

$$Precision = \frac{TP}{TP + FP}$$

Accuracy correct classification rate: This describes the total number of correct classifications with all classifications.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

Error rate, misclassification rate: This describes the total number of misclassifications about all classifications and is, therefore, the complement to accuracy.

$$Error\ rate = \frac{FP + FN}{TP + TN + FP + FN}$$

These key figures can be used to validate the selected NTS method and its application to the problem.

The following example should serve to illustrate the application of the confusion matrix. A method based on NTS data was created to determine surface water contamination with

wastewater. Using NTS data from contaminated and uncontaminated samples, the samples could grouped using a principal component analysis. The confusion matrix can be used to validate the grouping. For example, 12 contaminated samples (+) and 10 uncontaminated samples (-) are analyzed and evaluated. 10 of 12 contaminated samples were assigned correctly (+) and therefore 2 incorrectly (-). 9 out of 10 non-contaminated samples were correctly (-) assigned and thus 1 incorrectly (+). The results are entered in the confusion matrix in Figure J-5.

		Realität		
		positiv (A+)	negativ (A-)	Summe
Analyse	positiv (A+)	10 (TP)	1 (FP)	11 = TP + FP
	negativ (A-)	2 (FN)	9 (TN)	11 = FN + TN
	Summe	12 = TP + FN	10 = FP + TN	22 = PP +PN

Figure J-5: Confusion matrix of the sample data

This matrix can be used to calculate the characteristic data for validation:

Table J.1: Validation characteristics for the sample data

Sensitivity	0.83
Specificity	0.90
Precision	0.91
Accuracy	0.86
Error rate	0.14

With an accuracy of 86 %, the classification method used is very well suited for differentiating the samples. The error rate can be regarded as a measurement uncertainty for classification methods.

In addition to assessing the correctness and evaluation of a grouping of samples (classification), the confusion matrix can also be used to assess two (or more) different methods (evaluation/measurement) or laboratories for grouping samples. The confusion matrix is not limited to a binary classification but can be extended to several classes.

The following numerical example should serve to illustrate this. Based on the above example of grouping samples contaminated and uncontaminated with wastewater, two methods are available, each with an acceptable accuracy and error rate. For this purpose, 20 samples are

analyzed using both methods. The number of matches and differences in the classification by the two methods are noted in a confusion matrix (Figure J-6).

		Methode 2		
		positiv (A+)	negativ (A-)	Summe
Methode 1	positiv (A+)	9 (TP)	2 (FP)	11 (PP)
	negativ (A-)	1 (FN)	8 (TN)	9 (PN)
	Summe	10 (RP)	10 (RN)	20 (N)

Figure J-6: Confusion matrix of the sample data for the method comparison

Thus, 7 samples are classified as contaminated with Method 1 and Method 2, and 6 are classified as uncontaminated by both methods. Three samples are classified as contaminated with method 1 and uncontaminated with method 2, and 4 samples are classified as uncontaminated with method 1 but contaminated with method 2.

Cohen's kappa (κ) (kappa index) assesses the agreement between the two classification methods. *Cohen's Kappa (κ)* is a statistical measure for quantifying the degree of agreement between two assessments that classify objects (samples) into categories.

The calculation of κ is based on the following equation:

$$\kappa = \frac{p_o - p_e}{1 - p_e}$$

p_0: indicates the probability that both methods (LC-HRMS systems) provide the same result (o: observed).

$$p_o = \frac{TP + TN}{N}$$

p_e: indicates the theoretical probability of a random match (e: expected)

$$p_e = \frac{PP}{N} \cdot \frac{RP}{N} + \frac{PN}{N} \cdot \frac{RN}{N}$$

The evaluation of κ is based on the following table:

Table J.2: Assessment of agreement using the Kappa value

κ (Kappa)	**Compliance**
0.8 - 1.0	almost complete
>0.6	Strong
>0.4	Clear
>0.1	Weak
0 - 0.1	None

For the numerical example above. this results in κ = 0.7.
If the example data were a method comparison, for example, two LC-HRMS systems, there would be strong agreement.

Appendix K. Examples of quality controls

K.1 Example of quality control charts for a time-of-flight MS

QTOF measuring devices from Sciex have an automated injection system for automated mass calibration (so-called CDS - calibrant delivery system). The measurement files and results of this calibration are recorded approximately every 2 hours depending on the settings, and stored automatically. In addition to the raw spectrometric data, results data are also stored in text files, which are suitable for creating quality control charts.

An R-based application was created to read and analyze the stored data to generate an interactive quality control chart (Figure K-1). Mean values and distributions are calculated, and the results are displayed with definable warning limits[1]. Warning limits are system-dependent. The following warning limits are used in this example: Intensity (weighted[2] , averaged): Upper warning limit, lower warning limit; *m/z* deviation (averaged, ppm, mDa): Upper warning limit; Resolution (averaged): Lower warning limit.

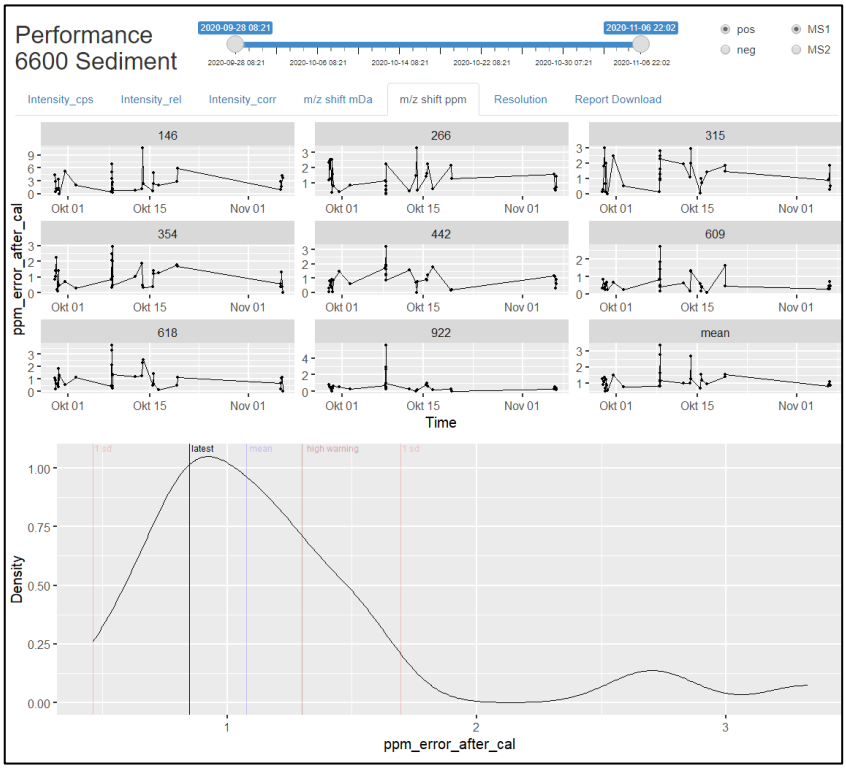


Figure K-1: Screenshot of an interactive quality control chart. Top line: Selection of period, polarity, and MS level. Second line (tab): Selection of parameters. Lines 3-5: Trend lines of the individual standards (named via their unit mass) and the mean value of all standards (mean). Sixth row: Density distribution of the mean value over the selected period. Vertical lines: Mean value of the last measurement ("latest"), total mean value, standard deviation, and warning limits defined by the user.

[1] The source code for this application is openly available (GPL 3) [https://github.com/bafg-bund/shewhart4cds]

[2] Weighting is necessary if individual standards have a high intensity compared to others and influence the mean value too strongly. The intensities of these are divided by a constant.

K.2 Example of quality control for an Orbitrap

High-resolution measuring devices from Thermo Fisher Scientific are regularly calibrated manually using the syringe pump and solutions provided by the device manufacturer. As commercial calibration solutions do not necessarily cover the desired mass range, these solutions could also be produced in-house.

The following parameters should be met for a successful calibration (LTQ Orbitrap, Classic): TIC variation < 5 % (positive) or < 10 % (negative), intensity of the most intense mass approx. 1E8, AGC target = 100 % and injection time < 2 ms (if the injection time is higher, the device should be cleaned).

The mass calibration is saved as a PDF; only the mass deviation is documented in ppm. If further information (TIC variation, intensity, injection time, etc.) is to be tracked, this must be noted or documented using a screenshot (Figure K-2). Automated documentation of the calibration data in the form of a control chart is currently unknown.

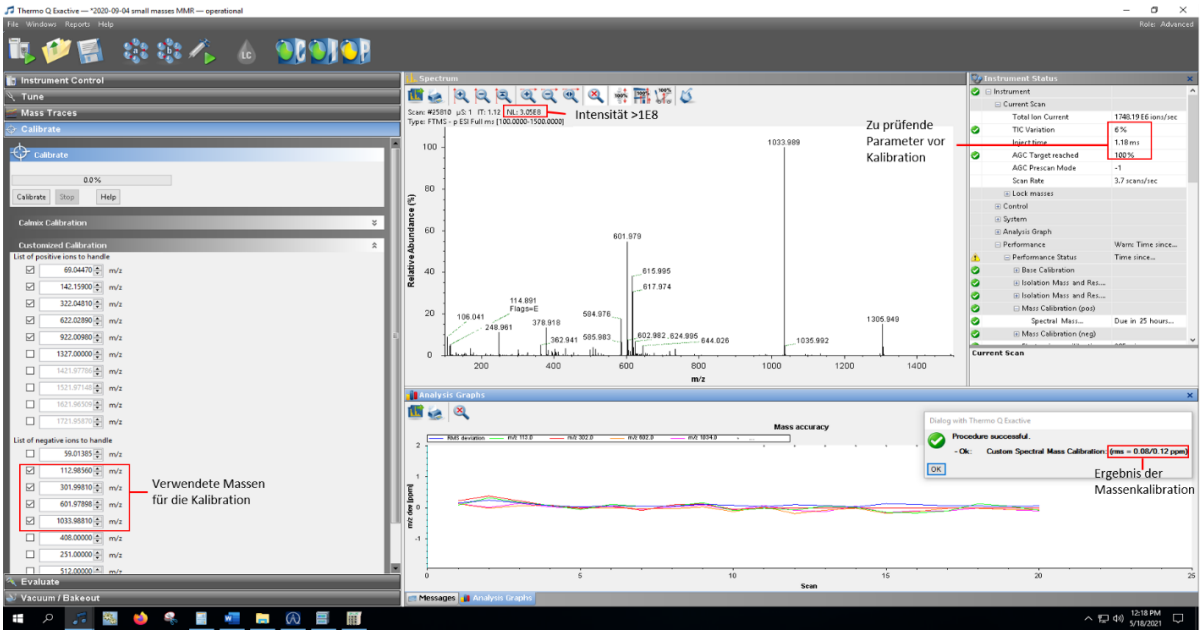


Figure K-2: *Screenshot of the tune page of an Orbitrap, on which the relevant parameters for quality control are visible (values marked in red).*

The duration for which the measuring device can be used after a calibration depends on many parameters (e.g., use of lock masses, temperature stability in the laboratory) and can be determined via the mass deviation ($\Delta m/z$) within a long sequence or confirmed via the evaluation function in the device's control software.

The system accepts an external calibration with a deviation of up to 5 ppm, which corresponds to the specification of the Orbitrap Classic. Experience has shown that the mass deviations after calibration are significantly more minor (< 2 ppm).