Joint Research Programme
BTO 2023.056 | Februari 2024

PFAS - Bedreiging voor de waterketen

Joint Research Programme

KWR

Bridging Science to Practice

CONCEPT

Report

PFAS - Bedreiging voor de waterketen

BTO 2023.056 | Februari 2024

This research is part of the Joint Research Programme of KWR, the water utilities and Vewin.

Project number
402045/153

Project manager
Patrick Bäuerlein

Client
BTO - Thematical research - Chemical safety

Author(s)
Elvio Amato, Dennis Vughs, Pascal Kooij, Frederic Béen

Quality Assurance
Thomas ter Laak

Sent to
This report is distributed to BTO-participants.

A year after publication it is public.

Keywords
PFAS, HRMS, analysis, suspect screening, NTS

Year of publishing
2024

More information
Dr Frederic Béen
T +31 30 606 9748
E frederic.been@kwrwater.nl

PO Box 1072
3430 BB Nieuwegein
The Netherlands

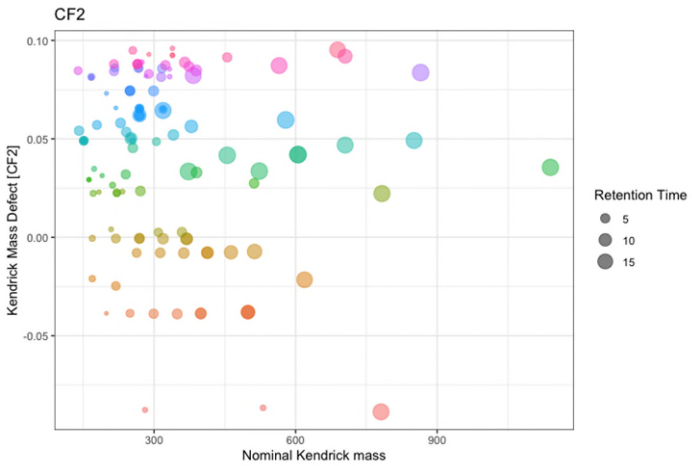
T +31 (0)30 60 69 511
F +31 (0)30 60 61 165
E info@kwrwater.nl
I www.kwrwater.nl

Managementsamenvatting

Inzetbare methoden voor gerichte en suspect en non-target screening (SNTS) van PFAS met ultrakorte tot lange fluorkoolstofketens

Authors Elvio Amato, Dennis Vughs, Pascal Kooij, Frederic Béen

In dit project zijn doelstof, suspect en non-target screeningmethoden (SNTS) ontwikkeld voor PFAS met ultrakorte tot lange fluorkoolstofketens, gebaseerd op hoge-resolutie-massaspectrometrie en de monstervoorbewerking met vaste-fase-extractie. De combinatie van gerichte en SNTS levert waardevolle informatie op over het voorkomen van deze PFAS in drinkwaterbronnen. SNTS is complex en kost tijd, maar levert aanvullende informatie op die niet kan worden verkregen met conventionele doelstofmethoden. De hier ontwikkelde workflow kan de verwerking van HRMS-gegevens voor PFAS-screening aanzienlijk versnellen en tegelijkertijd het aantal nieuwe verbindingen dat gedetecteerd kan worden verhogen. De aanpak kan worden gebruikt met gegevens verkregen met instrumenten die drinkwaterlaboratoria nu gebruiken en zowel historische gegevens (voor retrospectieve analyse) als nieuw geanalyseerde monsters verwerken. Zolang de gegevens worden verzameld in data-dependent acquisition (DDA) hoeven laboratoria hun huidige methoden niet aan te passen aangezien de workflow parallel kan worden toegepast op resultaten gegenereerd met de huidige instrumenten én de bestaande software voor gegevensanalyse.



Een Kendrick's Mass plot van potentiële non-target PFAS die in de monsters zijn gedetecteerd.

Belang: monitoring van PFAS in drinkwaterbronnen vereist uitgebreidere methoden

PFAS (per- and polyfluoroalkylverbindingen) vormen een brede en complexe groep chemische stoffen en vragen veel aandacht vanwege toenemende bezorgdheid over hun persistentie en potentiële toxiciteit. De huidige monitoringsmethoden richten zich op de analyse van een beperkt aantal relevante verbindingen uit PFAS-groep, maar omdat dit zo'n omvangrijke groep is, met zeer diverse eigenschappen, zijn brede methoden nodig om hun aanwezigheid en lot in het milieu in beeld te brengen.

Aanpak: gerichte en suspect en non-target screening op basis van massaspectrometrie

Om een uitgebreide methode voor de analyse van PFAS in drinkwaterbronnen te ontwikkelen, valideren en implementeren zijn doelstof- en suspect en non-target screeningmethoden (SNTS) ontwikkeld op

basis van hoge-resolutie-massaspectrometrie. De eerste stap bestond uit de validatie van een monster-voorbewerkingsmethode gebaseerd op vaste-fase-extractie. Vervolgens werden twee kwantitatieve methoden ontwikkeld en gevalideerd voor de analyse van een selectie van 36 PFAS met respectievelijk ultrakorte (< C4) en korte (≥ C4) tot lange (C12) fluor-koolstofketens. Deze methoden zijn vervolgens gebruikt om een SNTS te ontwikkelen. Na uitgebreide validatie zijn deze methoden gebruikt voor de analyse van zowel bronnen van drinkwater als geproduceerd drinkwater.

Resultaten: PFAS en precursoren ervan aanwezig en detecteerbaar in drinkwater(bronnen)

Uit gerichte analyses is gebleken dat 10 van de 36 gemeten verbindingen met hoge frequentie konden worden gedetecteerd in de meeste monsters van bronnen en geproduceerd drinkwater. Wat ultrakorte PFAS betreft, kon trifluorazijnzuur (TFA) op hoge niveaus (dus semi-kwantitatief) in bijna alle monsters worden gedetecteerd (30-370 ng/L). Bovendien werd TFA slechts in beperkte mate verwijderd; in sommige gevallen was de concentratie in het reinwater hoger dan in het ruwwater. Op basis van suspect en non-target screening konden andere PFAS worden geprioriteerd en voorlopig geïdentificeerd, waaronder 7 met relatief hoge betrouwbaarheid. Sommige van deze PFAS werden ook in gezuiverd drinkwater gedetecteerd. De resultaten van deze studie laten zien dat de combinatie van gerichte en SNTS waardevolle informatie oplevert over het voorkomen van PFAS in drinkwaterbronnen. Hoewel complex en tijdrovend, leveren SNTS aanvullende informatie op die niet kan worden verkregen met conventionele doelstof methoden. De methoden die in dit project zijn ontwikkeld, met name de SNTS workflow, kunnen in de toekomst worden gebruikt voor verkennende monitoringdoeleinden om potentiële nieuwe PFAS te detecteren die geen deel uitmaken van de routinemethode. Nieuwe stoffen die vaak worden gedetecteerd en/of een aanzienlijke intensiteit hebben ten opzichte van analoge legacy verbindingen (verbindingen die al langer voorkomen), kunnen formeel worden geïdentificeerd door referentiestandaarden aan te schaffen. Zo kunnen relevante nieuwe stoffen worden toegevoegd aan de lijst van stoffen die momenteel worden gemonitord.

Toepassing: gebruik de ontwikkelde workflow voor routine monitoring van PFAS

Drinkwaterlaboratoria hebben veel kennis over het gebruik en de toepassing van screeningsmethoden op basis van hoge-resolutie-massaspectrometrie (HRMS) en gebruiken die standaard als analyse-methoden om de kwaliteit van hun drinkwater-bronnen te onderzoeken. De grootste knelpunten bij HRMS-gegevens zijn echter de complexiteit en de lange verwerkingstijden. Met de hier ontwikkelde workflow kunnen laboratoria de verwerking van HRMS-gegevens voor PFAS-screening aanzienlijk versnellen en tegelijkertijd het aantal nieuwe verbindingen dat gedetecteerd kan worden verhogen. De aanpak kan tevens worden gebruikt voor gegevens verkregen met de instrumenten die drinkwaterlaboratoria al gebruiken en biedt zowel mogelijkheden voor retrospectieve analyse met historische gegevens als analyse van nieuwe monsters. Omdat de hier gebruikte software alleen gegevens uit data-dependent acquisition (DDA) ondersteunt, hoeven laboratoria hun huidige methoden niet aan te passen zolang ze dit gegevensformat gebruiken. De workflow kan parallel worden toegepast op resultaten gegenereerd met de huidige instrumenten én met de bestaande software voor gegevensanalyse. Drinkwaterbedrijven kunnen de ontwikkelde kennis over mogelijke aanwezigheid van PFAS die niet vallen onder de huidige monitoringprogramma's gebruiken om meer inzicht te krijgen in het type en de kenmerken van PFAS die in hun bronnen voorkomen of ze bedreigen.

Rapport

Dit onderzoek is beschreven in *PFAS - Bedreiging voor de waterketen* (BTO 2023.056).

Contents

Managementsamenvatting		**3**
Contents		**5**
1	Introduction	**6**
1.1	Background	6
2	Materials and methods	**7**
2.1	Chemicals	7
2.2	Sampling	7
2.3	Sample preparation	8
2.4	Method validation	9
2.5	Instrumental analysis	9
2.5.1	Target analysis	9
2.5.2	Suspect and non-target screening analysis	10
2.6	Data analysis	11
3	Results and discussion	**13**
3.1	Method validation and performance	13
3.1.1	C18 method (C ≥ 4)	13
3.1.2	MM method (C < 4)	13
3.2	Target analysis	13
3.2.1	Legacy PFAS	13
3.2.2	Ultra short-chain PFAS	15
3.2.3	Precursors and other fluorinated compounds	16
3.3	Suspect and non-target screening	17
3.3.1	Exploratory data analysis	19
3.3.1	Kendrick Mass Defect analysis	24
3.3.2	Raw versus treated drinking water feature profiles	26
3.4	Comparison between targeted and suspect and non-target screening	32
4	Conclusions	**33**
5	References	**34**
I	Appendix	**39**

1 Introduction

1.1 Background

Poly- and perfluoroalkyl substances (PFAS) are a broad group of anthropogenic organic substances which have been extensively used since the 1940s in numerous industrial processes (Ateia et al., 2019; Glüge et al., 2020). Due to their persistency, bioaccumulation, and toxicity, PFAS have been increasingly investigated around the world (Kaboré et al., 2018). In particular, these chemicals have been found in drinking water produced from sources in proximity of industrial sites, firefighting training areas and wastewater treatment plants (Eschauzier et al., 2013; Hu et al., 2016). This has led some C_8-C_{14} PFAS and their ammonium salts to be added to the candidate list of regulated substances in the Stockholm Convention (United Nations Environment Programme, 2019). More recently, the European Food Safety Authority (EFSA) proposed more stringent tolerable weekly intake (TWI) values (i.e., 4.4 ng/kg bw) for the sum of four common PFAS, namely perfluorooctanoic acid (PFOA), perfluorononanoic acid (PFNA), perfluorononanoic acid (PFHxS) and perfluorooctane sulfonic acid (PFOS) (Schrenk et al., 2020). Due to the restrictions imposed on PFAS, manufacturers are increasingly adopting alternative compounds, which often include short- or even ultra-short chain PFAS, as well as precursors and alternative (per)fluorinated compounds (e.g., perfluoropolyethers) (Kaboré et al., 2018; Z. Wang et al., 2013). The number of alternative fluorinated compounds currently in use has consequently been increasing and to date, almost 5000 chemicals have been classified as PFAS (OECD, 2023). However, when looking at repositories such as PubChem, the list of compounds which falls under the new PFAS definition increases up to over 7 million substances (E. Schymanski et al., 2023). The large number of compounds in use, combined with their (potential) toxicity and persistency, have given rise to concerns about the limitations of current analytical methods, in particular those used to monitor (drinking) water quality (Nakayama et al., 2019). In this respect, the European Commission adopted the new Drinking Water Directive (DWD) in which requirements for the development of analytical methods for the analysis of “PFAS total” and “Sum of PFAS” are stated (Drinking Water Directive, 2020), where the former refers to the analysis of “the totality of per- and polyfluoroalkyl substances” and the latter to the sum of 20 PFAS listed in point 3 of the DWD (Drinking Water Directive, 2020).

This study has two major goals, namely the (i) development and validation of targeted methods for the (semi-) quantitative assessment of a selection of PFAS and (ii) the development of a suspect and non-target screening (SNTS) strategy for the monitoring of a broader range of PFAS in drinking water sources and finished drinking water. For (semi-)quantitative analyses, two methods were developed and validated using conventional reversed-phase (RP) chromatography, for the analysis of short- to long-chain PFAS, and mixed-mode chromatography (MMC) for the semi-quantitative assessment of ultra-short PFAS. These same methods were then used in full scan mode to detect the potential presence of additional PFAS in samples from drinking water sources and finished drinking water. For this purpose, a workflow based on the recently developed R-package FluoroMatch (Koelmel et al., 2020, 2022) was implemented. The developed methods were then used to monitor PFAS in samples collected across different stages

Meer informatie
Dr Frederic Béen
T +31 30 606 9748
E frederic.been@kwrwater.nl

PO Box 1072
3430 BB Nieuwegein
The Netherlands

of the drinking water chain (i.e., groundwater, surface water, dune-filtrated water and drinking water) in the Netherlands and Belgium.

2 Materials and methods

2.1 Chemicals

All materials and equipment were rinsed with methanol (Merck, Darmstadt, Germany) and dried prior to use. Ammonium hydroxide solution (NH_4OH) was obtained from Sigma Aldrich. Ultrapure water (LiChrosolv®), LC-MS grade was obtained from Merck (Zwijndrecht, The Netherlands). Samples were analysed for 36 target PFAS and perfluorinated compounds including perfluoro-n-butanoic acid (PFBA), perfluoro-n-pentanoic acid (PFPeA), perfluoro-n-hexanoic acid (PFHxA), perfluoro-n-heptanoic acid (PFHpA), perfluoro-n-octanoic acid (PFOA), perfluoro-n-nonanoic acid (PFNA), perfluoro-n-decanoic acid (PFDA), perfluoro-n-undecanoic acid (PFUnDA), perfluoro-n-dodecanoic acid (PFDoDA), perfluoro-n-tridecanoic acid (PFTrDA), perfluoro-n-tetradecanoic acid (PFTeDA), 4:2 fluorotelomer sulfonic acid (4:2 FTS), 6:2 fluorotelomer sulfonic acid (6:2 FTS), 8:2 fluorotelomer sulfonic acid (8:2 FTS), hexafluoropropylene oxide dimer acid (HFPO-DA), 4,8-dioxa-3H-perfluorononanoic acid (NaDONA), 9-chlorohexadecafluoro-3-oxanonane-1-sulfonic acid (9Cl-PF3ONS), 11-chloroeicosafluoro-3-oxaundecane-1-sulfonic acid (11Cl-PF3OUdS), perfluoro-n-butanesulfonic acid (PFBS), perfluoropentane sulfonic acid (PFPeS), perfluoro-n-hexanesulfonic acid (PFHxS), perfluoro-n-heptanesulfonic acid (PFHpS), perfluoro-n-octanesulfonic acid (PFOS), perfluoro-n-nonane sulfonic acid (PFNS), perfluoro-n-decanesulfonic acid (PFDS), perfluoro-1-butanesulfonamide (FBSA), perfluoro-1-hexanesulfonamide (FHxSA), perfluoro-1-octanesulfonamide (FOSA), N-methylperfluorooctanesulfonamidoacetic acid (N-MeFOSAA), N-ethylperfluorooctanesulfonamidoacetic acid (N-EtFOSAA), trifluoromethanesulfonamide (PFMeSAm), pentafluoroethanesulfonamide (PFEtSAm), trifluoroacetic acid (TFA), trifluoromethanesulfonic acid (F3-MSA), perfluoropropanoic acid (PFPrA), and perfluoroethane sulfonic acid (PFEtSA). Mass-labelled internal standard (IS) consisting of a mixture of 19 carboxylic and sulphonic PFAS were purchased from Wellington Laboratories (Ontario, Canada). The IS HFPO-DA-$^{13}C_3$ was supplied by Greyhound Chromatography and Allied Chemicals (Birkenhead, United Kingdom).

2.2 Sampling

Samples were provided by 11 drinking water companies (10 from the Netherlands and 1 from Belgium) and consisted of surface water, groundwater, dune and riverbank infiltrated water, and drinking water (see Table 1). It should be noted that samples from location S11 were collected at a later stage and were not analysed in the same sequence as the other samples. Hence, only quantitative (targeted) results are reported for S11 given that for SNTS, sample should have been analysed in a single sequence to minimise instrumental variability. Samples were collected using polypropylene (PP) bottles cleaned with MeOH and stored at -20°C until analysis. For method validation, drinking

water consisted of tap water collected from our laboratory, while surface water was obtained from a canal in Nieuwegein (the Netherlands).

Table 1: Summary of sampling locations. These locations were also included in the recent study by (Sadia et al., 2023). Information about the identity of each location has been provided separately to this report.

Location	Type of water sampled	Code	Sampling date
S1	Surface water (SW)	S1-SW	04-06-2021
	Finished drinking water (DW)	S1-DW	04-06-2021
S2	Surface water (SW (a))	S2-SW (a)	07-06-2021
	After storage basin – Surface water (SW (b))	S2-SW (b)	07-06-2021
	Drinking water (DW)	S2-DW	07-06-2021
S3	Surface water (SW)	S3-SW	07-06-2021
	Finished drinking water (DW)	S3-DW	07-06-2021
S4	Surface water (SW)	S4-SW	07-06-2021
	Finished drinking water (DW)	S4-DW	07-06-2021
S5	After storage basin – Surface water (SW)	S5-SW	10-06-2021
	Finished drinking water (DW)	S5-DW	10-06-2021
S6	Groundwater (GW)	S6-GW	07-06-2021
	Drinking water (DW)	S6-DW	07-06-2021
S7	Groundwater (GW)	S7-GW	07-03-2021
	Drinking water (DW)	S7-DW	07-03-2021
S8	Groundwater (GW)	S8-GW	17-06-2021
	Drinking water (DW)	S8-DW	17-06-2021
S9	Surface water (SW (a))	S9-SW (a)	07-06-2021
	After storage basin – Surface water (SW (b))	S9-SW (b)	07-06-2021
	After dune filtration (DF)	S9-DF	07-06-2021
	Drinking water (DW)	S9-DW	07-06-2021
S10	After riverbank filtration (DF)	S10-DF	07-06-2021
	Drinking water (DW)	S10-DW	07-06-2021
S11	Surface water (SW (a))	S11-SW (a)	08-12-2021
	Surface water (SW (b))	S11-SW (b)	08-12-2021
	Drinking water (DW)	S11-DW	08-12-2021

2.3 Sample preparation

PFAS and other perfluorinated compounds were extracted by means of solid phase extraction (SPE). Water samples collected in the field were allowed to defrost overnight before extraction. Subsamples of 500 mL were transferred into precleaned plastic bottles and procedural laboratory blank consisting of ultrapure water were included in the extraction batch. Following addition of internal standard (100 µL mixture of 20 mass-labelled PFAS (20 ug/L in MeOH) and 100 µL of HFPO-DA ($^{13}C_3$, 20 µg/L in MeOH)), samples were acidified to a pH of 4-5 using HCl. The SPE was carried out using Oasis-WAX SPE cartridges (150 mg, 30 µm, 6cc, Oasis®). The extraction procedure included a precleaning step which was performed using a polypropylene cartridge containing 3 grams of sea sand placed on top of the Oasis-WAX cartridge using an adapter. Before extraction, both cartridges were conditioned using two times 5 mL 0.25% ammonium hydroxide (NH_4OH, 29%) in MeOH followed by two times 5 mL ultrapure water (LC-MS grade). Sand

cartridges and adapters were removed and 5 mL of 25 mM ammonium acetate pH 4 was added to the WAX SPE cartridges for cleaning. SPE cartridges were covered with aluminium foil and left to dry for 1.5 hours under vacuum.

Empty bottles were allowed to dry and rinsed using 10 mL of MeOH. The rinsing solution was transferred to a clean glass tube. The same tube was then used to collect the eluate from the corresponding SPE cartridges. Cartridges were eluted using 5 mL 0.25% ammonium hydroxide (NH_4OH, 29% in MeOH). Eluates (i.e., 15 mL in total) were reduced to a volume of 500 µL under a gentle stream of nitrogen (Barkey, 67°C). After evaporation, samples were made up to a final volume of 1 mL (MeOH:ultra-pure (UP) water 1:1 v/v) spiked with an injection standard (IS) PFAS mixture (i.e., $^{13}C_3$PFBA, $^{13}C_2$PFOA and $^{13}C_4$PFOS (4 µg/L in MeOH)), resulting in a concentration factor of 500x. Samples were then filtered (0.45 µm, regenerated cellulose, Whatman®) and stored in the freezer (-20 °C) until analysis. Procedural blanks were prepared using the same bottles used for sample collection (i.e., from the same batch), filled with ultrapure water and extracted using the procedure described above.

2.4 Method validation

Aliquots of 500 mL of drinking and surface water were spiked to final concentrations of 0.2, 1 and 5 ng/L for PFAS analytes and 4 ng/L for IS. For each matrix (i.e., drinking and surface water), five replicates and one blank were prepared. These were extracted and analysed by LC-HRMS in duplicate. Limits of detection (LOD) were reported as 3 times the standard deviation of the lowest spike level with a relative standard deviation (RSD) < 10%. When RSD values of all spike levels were > 10% the standard deviation with the lowest RSD was used. Limits of quantification (LOQ) were reported as 3 times the LOD. Matrix effects were determined using water samples spiked with PFAS at final concentration levels of 0.4, 0.8, 1.2, 1.6 and 2 µg/L. Samples were extracted following the procedure described previously, and calibration curves were obtained by plotting concentrations measured in each level against the corresponding instrumental response (i.e., peak areas). Matrix effects were estimated as the ratio between the slope of the matrix curve and the slope of a calibration curve obtained using pure standards, expressed as a percentage. Recoveries were calculated by dividing the concentration measured in spiked samples (after blank correction) by the nominal spiked concentration, reported as a percentage. Finally, Intra-day instrument accuracy and precision was determined by eight quantifications of a 2 µg/L standard and reported as RSD.

2.5 Instrumental analysis

2.5.1 Target analysis

Analyses were performed using a Tribrid Orbitrap Fusion mass spectrometer (ThermoFisher Scientific, Bremen, Germany) equipped with a heated electrospray ionization source interfaced to a Vanquish HPLC system (ThermoFisher Scientific). Two chromatographic methods, i.e., Reverse Phase (RP) and Mixed Mode (MM) were used. RP chromatographic analysis was performed using an XBridge BEH C18 XP column (100 mm × 2.1 mm I.D., particle size 2.5 µm, Waters, Etten-Leur, The Netherlands) in combination with a 2.0 mm × 2.1 mm I.D. Phenomenex SecurityGuard Ultra column (Phenomenex, Torrance, USA), maintained at 25°C. Eluent composition was A: 2 mM NH4Ac in ultrapure-water and B: 2 mM NH4Ac in MeOH:ACN 80:20% (v/v). The LC gradient started with 30% B for 1

min. Then increased to 60% B in 2 min and stayed constant for 4 min. Next, increased to 90% B in 9 min and then increased to 100% B in 1 min and stayed constant for 3 min. Due to lacking retention of six ultra-short chain (C < 4) PFAS (TFA, PFPrA, F3-MSA, PFEtS, PFMeSAm and PFEtSAm) on reversed-phase columns, an alternative semi-quantitative LC method was developed and optimized for these ultra-short chain compounds using an Trinity P1 Mixed Mode column (50 mm × 2.1 mm I.D., particle size 3 µm, ThermoFisher). The Trinity P1 column was kept at 35°C during analysis. For MM analysis, eluent A consisted of ultrapure water with 2 mM ammonium acetate at a pH of 4.5. Eluent B consisted of 80% acetonitrile and 20% ultrapure water (v/v) with 40 mM ammonium acetate at a pH of 4.5. The gradient started constant with 10% B for 1 min. Then increased linear to 60% B in 5 min, and then increased to 100% B in 7 min and stayed at 100% B for 11 min. The injection volume and the flow rate were 10 µL and 0.3 mL/min for both methods, respectively.

Mass spectrometric detection was performed in negative ionization mode for RP and MM. The Vaporizer and capillary temperature was set to 300 °C and 200 °C, respectively. Sheath, auxiliary and sweep gas were set to arbitrary units of 30, 5 and 5 for both RP and MM. The source voltage was set to − 2.5 kV in negative mode. The RF lens was set to 50% for RP and 30% for MM. Full scan high resolution mass spectra were recorded from m/z 150–1000 for RP and 70-500 m/z for MM, with a resolution of 120'000 FWHM. Quadrupole isolation was used for acquisition with a 5 ppm mass window. The AGC target was set to 200'000 and maximum injection time was set to 100 ms. MS/MS spectra were recorded using a mass table list. Optimal HCD collision energy was selected for each individual PFAS. With every batch run, mass calibration was performed using a Pierce FlexMix calibration solution to obtain a mass error of < 2 ppm.

Calibration standards for both methods were prepared in 50:50 MeOH:H2O (v/v) at concentration levels of 0, 0.25, 0.50, 1, 2.5, 10, 25, 50 µg/L. PFAS injection standards consisting of $^{13}C_3$PFBA, $^{13}C_2$PFOA and $^{13}C_4$PFOS (4 µg/L in MeOH) were used as instrument performance standards. The calibration curve was analysed before the sample sequence, and blank runs were collected before and after calibration sequences and in between samples of different compounds.

2.5.2 Suspect and non-target screening analysis

Suspect and non-target screening analyses were performed using the same LC-HRMS system and chromatographic conditions as described above. Only the gradient for the RP method was slightly adjusted (i.e., 5 min of 100% B instead of 3 min of 100% B to include larger PFAS molecules that elute later and may be missed with the targeted method). For SNTS some adjustments were made to the mass spectrometric method. The scan range for RP method was increased to 120 – 1500 m/z, while the scan range for MM was not adjusted. Instead of a mass list table, data-dependent acquisition (DDA) was performed using a high-collision dissociation (HCD) energy at 20, 35 and 50% (stepped) and a FT resolution of 15 000 (FWHM).

2.6 Data analysis

Peak picking, componentization and preliminary suspect screening of NTS data for both C18 and MM were performed using Compound Discoverer 3.2 (ThermoFischer Scientific). Chromatograms were aligned between 1.5-22min and 1.5-20min for C18 and MM, respectively. A mass tolerance of 3ppm was used for feature grouping, molecular formula prediction and isotope search. Library searches were performed using a mass tolerance of 3 ppm and a maximum retention time shift of 0.5 min. Searches were performed against *ChemSpider* (Royal Society of Chemistry), *mzCloud* (HighChem Ltd., Slovakia) and *mzVault* (ThermoFischer Scientific), as well as a set of PFAS available (i.e., OECDPFAS (Z. Wang et al., 2018), PFASNTREV19 (NORMAN Network, 2023)) and in-house curated libraries. After processing, the generated feature list was exported to .csv for further analysis using RStudio (RStudio Team, 2021). In parallel, results were also processed using FluoroMatch Modular (Koelmel et al., 2020, 2022). In particular, .raw files were initially converted to .ms2 files using MSConverter and then the FluoroMatch Modular workflow was used to process a selection of samples. Following the proposed approach in FluoroMatch, "pooled" samples are used for actual peak picking and feature annotation, in combination with a feature list obtained with another processing software (in this case the above-mentioned feature list generated with Compound Discoverer was used). Because pooled samples were not used in this study, the FluoroMatch Modular workflow was implemented on one sample per location to cover the whole range of expected PFAS types and concentrations. FluoroMatch Modular parameters were set according to the standard settings suggested by the authors for Orbitrap instruments and the specific acquisition settings used for the DDA experiments (Koelmel et al., 2020, 2022).

The newly generated feature list containing information from both FluoroMatch and Compound Discoverer was imported into RStudio for further processing. Features which did not show an intensity (i.e., average peak height of triplicate injections) of at least 10x that measured in blanks or which were not above 4e5 counts (i.e., noise threshold) were removed. Similarly, features showing a coefficient of variation between injections above 15% were removed. Subsequently, principal component and hierarchical cluster analysis (PCA and HCA, respectively) were used to investigate the presence of specific sample groups and, in particular, prioritise (abundant) PFAS-like features for further analysis. PCA was applied to all samples after scaling feature intensities (mean of triplicate injection) using the *prcomp* function present in the *stats* package. Prior to HCA, group mean feature intensities were normalised by the maximum. The Euclidean distance and Ward's clustering method were used. HCA was used on all samples together as well as on samples from individual locations. The latter was used to prioritise PFAS-like features which appear not to be removed during treatment or features whose intensity increases either after treatment (e.g., transformation products) or which appear to increase after specific treatments (e.g., contribution of sea-spray aerosols during dune infiltration). HCA and corresponding heatmaps were computed using the *pheatmap* function from the *pheatmap* package in R. Volcano plots were computed using log2 fold-changes and p-values obtained from a Student's t-test and corrected using the Benjamini & Hochberg method. Volcano plots were specifically used to prioritise features which increased after treatment in cases where samples were collected only at two stages of the treatment process (raw vs treated water). To further refine the prioritisation process, the scoring system included in FluoroMatch, ranging from A (confident tentative identification based on exact mass, in-silico fragmentation pattern and/or match with a spectral library, equivalent to level 2 or 3 according to Schymanski et al.(2014)) to E (likely not a

PFAS feature), was considered during data visualisation. Detection of homologous series based on Kendrick Mass Defect (KMD) analysis was performed following similar strategies as described in (Bugsel & Zwiener, 2020; Jacob et al., 2021; Koelmel et al., 2020). Briefly, FluoroMatch outputs, which already include information about the appurtenance of detected features to specific homologous series (HS) were further processed. Focus was set on HS with the following repeating units: CF2, CF2CH2, OCF2CFCF3 and CF2CFCl. For the first two repeating units, a minimum of 3 features per HS was used as a threshold for further prioritisation. For the other two units, a minimum of 2 features per HS was used as it was noted that 3 would be too stringent and no feature would remain for further analysis. In addition, the focus was set only on features having a KMD between -0.25 and 0.1 (Bugsel & Zwiener, 2020), to reduce the number of features to be analysed. Features were also further prioritised based on their retention time, which needed to be coherent with the increasing number of carbon atoms in the annotated formula. Detailed information about the mass algorithm and mass tolerances used in FluoroMatch can be found in (Koelmel et al., 2020, 2022).

Annotated structures proposed by FluoroMatch were further verified using the databases for suspect screening mentioned above, as well as MassBank EU (Schulze et al., 2021). In case the annotated structure was not available in existing spectral libraries, *in-silico* fragmentation of the proposed structure was performed using *MetFrag* and/or the Fragment Ion Search (FISh) Scoring tool included in Compound Discoverer. (Tentative) Identification confidence of prioritised features was then expressed using the scale recently proposed by Charbonnet et al. (Charbonnet et al., 2022). In brief, this scale, which ranges from 1a (formal identification with a reference standard) to 5b (non-target PFAS exact mass), is based on previous work by (E. L. Schymanski et al., 2014) with the addition PFAS specific information, such as whether the feature is part of a homologous series, whether it's a potential positional isomer and so forth.

3 Results and discussion

3.1 Method validation and performance

3.1.1 C18 method (C ≥ 4)

For drinking water samples, LOD and LOQ ranged between 0.02 and 0.84 ng/L, and between 0.06 and 2.52 ng/L, respectively (Appendix I.III). Recoveries were consistently between 88 and 108%, except for FHxSA, FBSA, 11Cl-PF3OUdS and PFTrDA, which had recoveries of 66, 60, 57 and 48%, respectively (Appendix I.III). For surface water samples, LOD and LOQ ranged between 0.03 and 0.39 ng/L, and between 0.09 and 1.16 ng/L, respectively, and recoveries ranged between 73 and 130%. For both drinking and surface water analyses, reporting limits were set to 0.5 ng/L, except for PFTrDA and PFTeDA, which had a reporting limit of 3 ng/L (Appendix I.III). Results for the chromatographic and mass spectrometry optimization are shown in Appendix I.VII and I.VIII, respectively.

3.1.2 MM method (C < 4)

LOD and LOQ estimated for drinking water analyses ranged between 0.07 and 1.79 ng/L, and between 0.22 and 5.36 ng/L, respectively (Appendix I.IV). Similarly, LOD and LOQ measured for surface water ranged between 0.08 and 1.79 ng/L, and between 0.24 and 5.36 ng/L, respectively (Appendix I.IV). Recoveries ranged from 15 to 102%, for drinking water, and from 14 to 210%, for surface water, respectively. Reporting limits were the same for both drinking and surface water sample and ranged between 1 and 5 ng/L, except for PFPrA and TFA which had a higher reporting limit, i.e., 10 and 20 ng/L, respectively. Results for the chromatographic and mass spectrometry optimization are reported in Appendix I.VII and I.VIII, respectively.

3.2 Target analysis

3.2.1 Legacy PFAS

The DWD provides environmental quality standards for a total of 20 PFAS, which are characterised by a number of carbons ranging from 4 to 13 atoms. In this study, 17 out of the 20 PFAS listed in the DWD were measured in water samples, including an additional PFAS with 14 carbon atoms (PFTeDA). Among these, the most frequently detected compound was PFBA, which occurred in all 23 samples, followed by PFOA, PFPeA, PFHxA, PFBS, PFHpA, PFOS, PFHxS, PFNA, and PFDA, which were found in 20, 18, 17, 17, 15, 12, 9, 3 and 1 sample, respectively (Figure 1 and Figure 2). HFPO-DA was detected in 3 SW samples (i.e., 0.64, 20 and 20 ng/L for S3, S5 and S10, respectively), and 2 DW samples (0.59, and 9.0 ng/L for S5 and S10, respectively). The remaining PFAS, i.e., PFDoDA, PFDS, PFHpS, PFNS, PFPeS, PFTeDA, PFTrDA, and PFUnDA, occurred at concentrations below the LOQ of the method (Appendix I.III). Total PFAS concentrations measured in drinking water samples were consistently lower than those measured in the correspondent water sources (i.e., groundwater, surface water, and dune-filtered water), and ranged between 0.78 and 29 ng/L (Figure 1 and Figure 2). On average, drinking water concentrations were 54% lower than those measured in water sources. Increasing total PFAS concentrations were observed following dune filtration in S9-DF, suggesting

that dunes may be a potential source of PFAS contamination. Among the 20 PFAS listed in the DWD, PFHxS, PFOS, PFOA and PFNA, are subject to stricter regulations in the European Union due to their greater toxicity, i.e., the European Food Safety Authority (EFSA) has set a maximum combined intake for these four compounds of 4.4 ng/kg body weight per week (Schrenk et al., 2020). This threshold value roughly corresponds to a water concentration of approximately 4.4 ng PFOA equivalents (PEQ)/L (van der Aa & te Biesebeek, 2021), which is considerably lower than the EQS value for the total PFAS concentration in water of 500 ng/L (Proposal for a Directive of the European Parliament and of the Council on the Quality of Water Intended for Human Consumption, 2020). PFOA equivalents can be used to estimate the expected toxicity of mixtures of PFAS. This approach involves multiplying the concentration of a given compound by its potency factor relative to PFOA (RPF), and adding up the PEQ concentration of each individual PFAS. Thus, for the 4 PFAS that are more strictly regulated by EFSA (i.e., EFSA-4), the PEQ concentration is obtained using

$$\mathrm{EFSA4} = ([\mathrm{PFOS}]_{\mathrm{L+B}} * \mathrm{RPF}_{\mathrm{PFOS}}) + ([\mathrm{PFOA}]_{\mathrm{L+B}} * \mathrm{RPF}_{\mathrm{PFOA}}) + ([\mathrm{PFHxS}]_{\mathrm{L+B}} * \mathrm{RPF}_{\mathrm{PFHxS}}) + ([\mathrm{PFNA}]_{\mathrm{L+B}} * \mathrm{RPF}_{\mathrm{PFNA}})$$

where RPF of 2, 1, 0.6 and 10 are used for PFOS, PFOA, PFHxS and PFNA, respectively. L and B refer to linear and branched molecules (though only linear molecules were determined in this study). The combined PEQ concentration of the four compounds in drinking water was consistently <4.4 ng PEQ/L, except for S5-DW and S9-DW which slightly exceeded this threshold (i.e., 5.7 and 5.3 ng PEQ/L, respectively) (Figure 3). In general, PFOA was the most abundant and frequently detected among the four PFAS. In water sources, PFOA concentrations always exceeded those of PFHxS, PFOS, and PFNA, in some cases of a factor up to 10. In drinking water, PFOA was generally the only compound detected, suggesting that PFOA removal during water purification treatments may still pose some challenges.

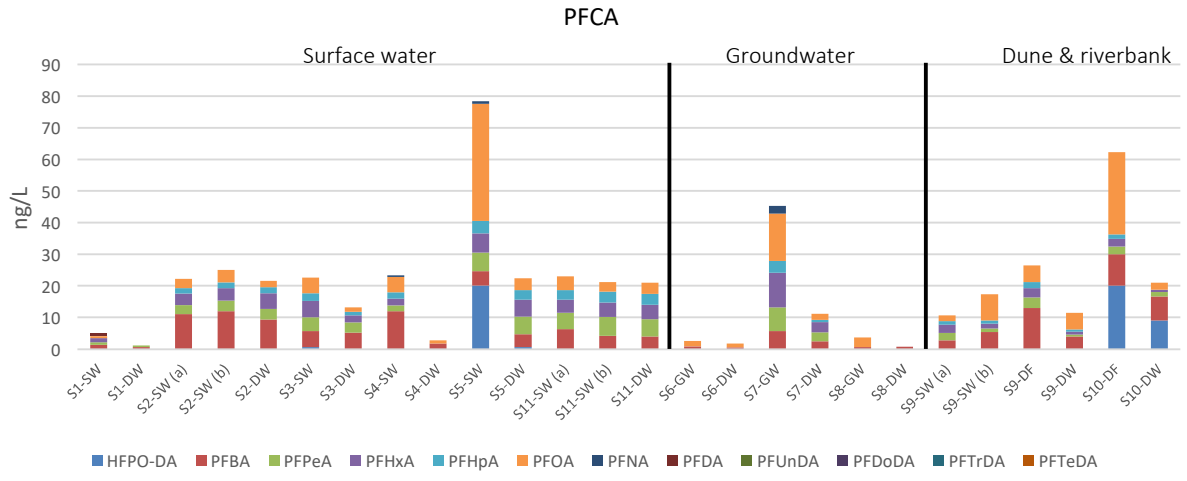


Figure 1: Detected legacy carboxylic acids (PFCA) listed in the DWD sorted by type of sample (i.e., surface and groundwater, dune and riverbank infiltrate).

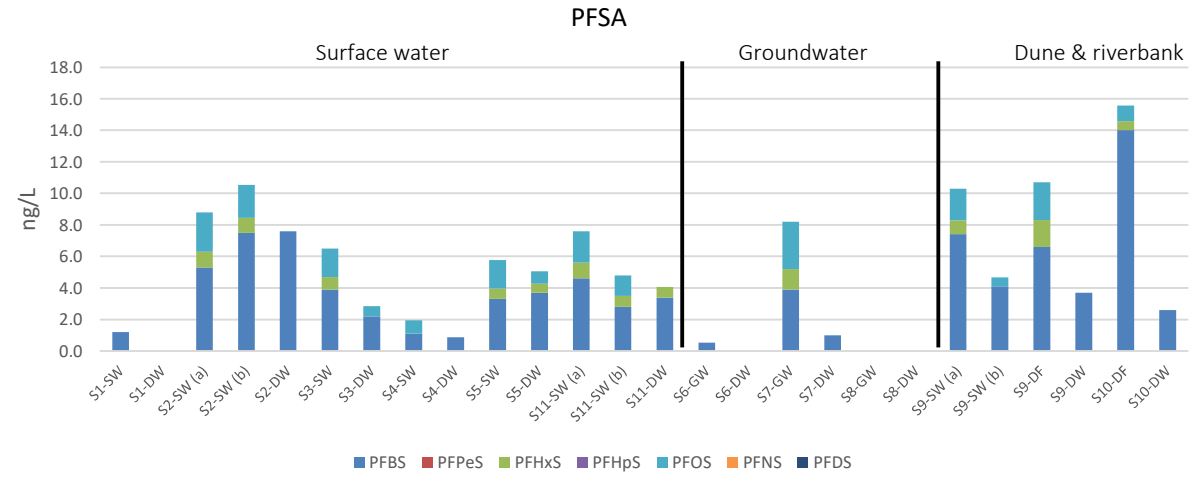


Figure 2: Detected legacy sulfonic acids (PFSA) listed in the DWD sorted by type of sample (i.e., surface and groundwater, dune and riverbank infiltrate).

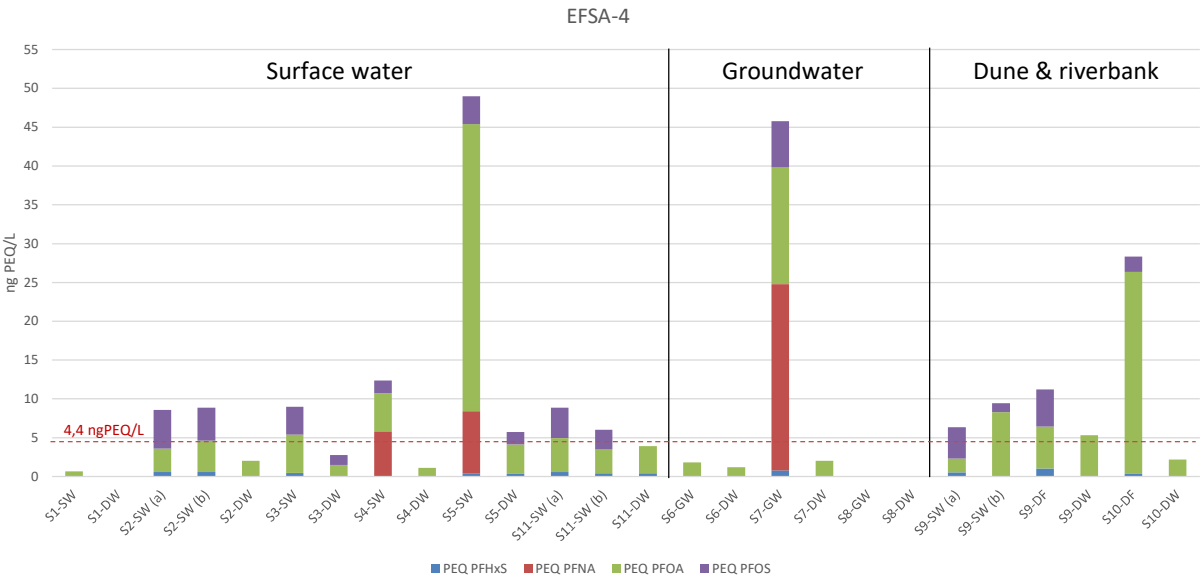


Figure 3: occurrence of EFSA-4 in water samples sorted by type of water.

3.2.2 Ultra short-chain PFAS

Short-chain PFAS (i.e., ≤3 carbon atoms) were measured only semi-quantitatively as no internal standards were available to account for extraction recoveries. Hence, reported results should be considered only as indicative. Overall, ultra short-chain PFAS were detected at much higher concentrations compared to short- and long-chain PFAS (Figure 4). TFA was by far the most abundant compound, occurring at concentrations ranging between < LOQ and 800 ng/L, in water sources, and between, < LOQ and 705 ng/L, in drinking water, respectively. F3-MSA concentrations ranged between < LOQ and 55 ng/L in water sources, and between < LOQ and 31 in drinking water, while PFPrA occurred in two water sources at 11-18 ng/L and in two samples of drinking water samples at 11-12 ng/L, respectively. TFA was also the most frequently detected compound (i.e., 24 samples), followed by F3-MSA (19 samples) and PFPrA (6 samples). PFEtSA was always detected at concentrations below the LOQ (Appendix I.IV). Similarly to the trend observed for their analogous medium- and long-chain, also ultrashort-chain PFAS appeared to show lower total

concentrations in drinking water samples compared to source water samples, except for S2-DW, S3-DW, and S5-DW samples. Also, TFA and F3-MSA showed a considerable increase in concentrations following dune filtration in S9 samples (from 50 to 285 ng/L, and from 3.5 to 17 ng/L, for TFA and F3-MSA, respectively), as previously observed for some medium- and long-chain PFAS. The overall higher concentrations of short-chain (especially TFA) compared to medium- and long-chain PFAS seems to indicate that these compounds may occur more frequently and at higher concentrations in both source and drinking water. Higher concentrations of short-chain PFAS in water may be expected due to their strong hydrophilic behaviour and tendency to partition to the aqueous phase, which makes them more mobile and more difficult to remove during water purification processes compared to long- and medium-chain PFAS (Eschauzier et al. 2012; Zhao et al. 2016; Reemtsma et al. 2016; Ateia et al. 2019; Gagliano et al. 2020). Furthermore, these PFAS might also be transformation products of larger PFAS which are broken down during treatment (Scheurer et al., 2017). This could be the case for locations in which an increase in for instance TFA concentrations were observed after treatment. While the analytical method used for short-chain PFAS is semi-quantitative, and care should be taken when interpreting these results, data indicate that also these compounds should also be included in monitoring programmes and their potential toxic effect adequately assessed (Nian et al. 2020).

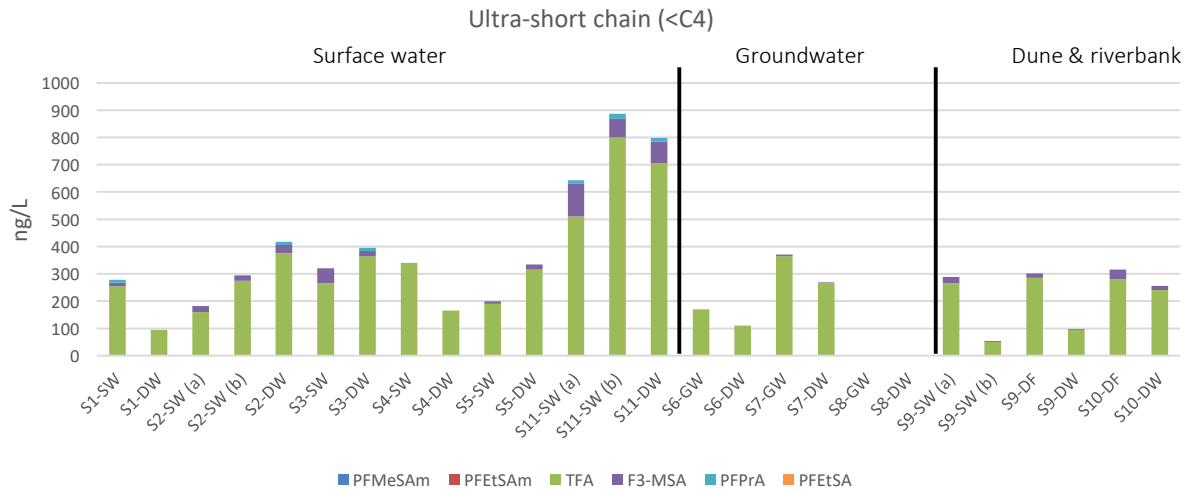


Figure 4: occurrence of short-chain PFAS in water samples grouped by type of water. ***NB results reported here are semi-quantitative and are hence only indicative.***

3.2.3 Precursors and other fluorinated compounds

Concentrations of PFAS precursors (i.e., 4:2 FTS, 6:2 FTS, 8:2 FTS, FBSA FHxSA, FOSA, NaDONA, N-MeFOSAA, PFMeSAm, and PFEtSAm) measured in water samples were generally below the LOQ, except for 6:2 FTS and FBSA, which were found in 3 and 1 sample, respectively (between 0.63 and 1.5 ng/L), as shown in Figure 5. Also

concentrations of other perfluorinated compounds (i.e., 11Cl-PF3OUdS, 9Cl-PF3ONS, and HFPO-DA) generally occurred at concentrations below the LOQ.

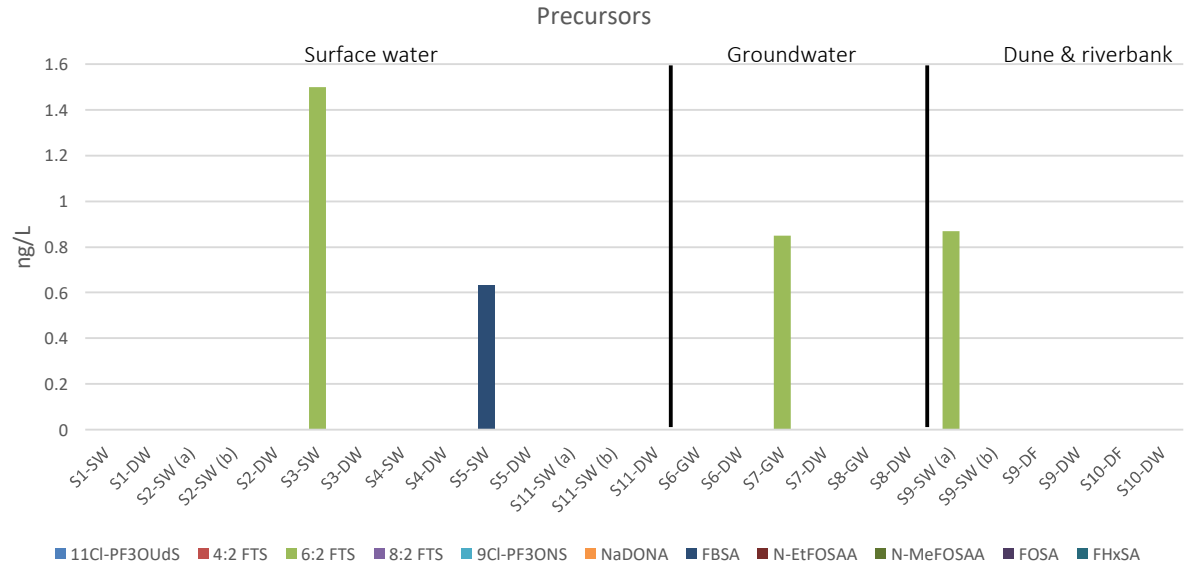


Figure 5: Precursor concentrations [ng/L] measured in the collected samples.

3.3 Suspect and non-target screening

Following peak picking, componentisation and feature annotation using Compound Discoverer and FluoroMatch Modular, a total of 10826 features were initially detected. As described previously, retention time, background signal and group coefficient of variation (CV) were used to filter out non-relevant features, resulting in 8 157 features which were retained for further processing. Out of features detected using the C18 method, 15 were scored A, 1 764 B (of which 8 B+), 2 988 C, 1253 D and 2657 E, according to the scoring system included in FluoroMatch. In the case of features detected with the MM method, 4 were labelled A, 10 B (of which 2 B+), 8 C, 380 D and 453 E. Following a similar strategy as proposed by Jeong et al. (Jeong et al., 2022), features labelled as E were excluded a priori from further analyses because these are likely not PFAS compounds (Koelmel et al., 2020).

FluoroMatch workflow and scoring system

FluoroMatch is a dedicated data analysis workflow which has been recently developed by Koelmel and colleagues (Koelmel et al., 2020, 2022). The workflow functions in an analogue way to conventional vendor (e.g., Compound Discoverer, MetaboScape, MassHunter) or open source (e.g., patRoon, MSDial) software used to process HRMS data (e.g., peak picking, deconvolution and alignment of features) yet it differs in the fact that it was specifically developed for PFAS analysis. In particular, the software uses large PFAS libraries (e.g., US EPA) as well as *in silico* fragmentation data specific for PFAS.

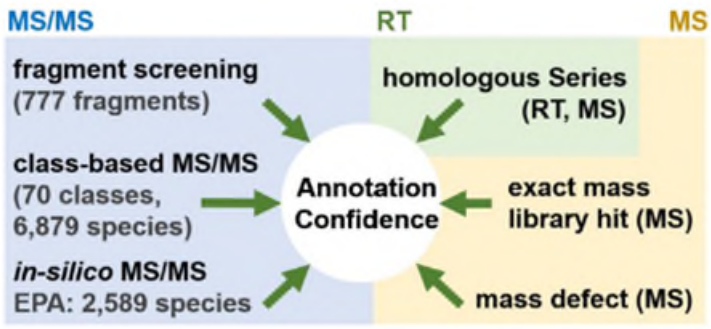


Figure 6: PFAS specific MS information used by FluoroMatch to annotate detected features.

This means that during the annotation step (i.e., trying to attribute a structure to a detected feature), it will focus only on features that could potentially be a PFAS compound (i.e., likely containing fluorine in their molecular formula). Furthermore, FluoroMatch also automatically provides information on whether detected features belong to so called homologous series (more information on this is provided in 3.3.1), consequently increasing the identification confidence. Once all this information is combined, FluoroMatch will also attribute a score, going from A (confident identification) to E (likely not a PFAS), to each of the detected features (see Figure 7). This score gives an indication about the identification confidence and is inspired by the confidence levels suggested by Schymanski et al. (2014) and Charbonnet et al. (2022). As mentioned earlier, the final confidence in (tentative) identification of features detected in this work was expressed using the scale recently proposed by Charbonnet et al. (2022) as it directly relates to the commonly used scale by Schymanski et al. (2014) (see Figure 8).

Score	Description	Score	Description
A+	Confirmed by standards (RT, MS/MS, MS)	C+	In series w/ confident ID
A	Confident ID (MS/MS, MS) and series	C	In series (3+) w/ any tentative ID
A-	Confident ID (MS/MS, MS)	C-	In series w/ any tentative ID
B+	Tentative ID (MS/MS, MS) and series	D+	MD or MS and in series (3+)
B	Tentative ID (MS/MS, MS) and series	D	In series (3+)
B- *B--	Any tentative ID (MS/MS, MS)	D-	MD or MS

Figure 7: Scoring system used in FluoroMatch to express the identification confidence. "and series" means that a feature is part of a homologous series.

Level	Identification Confidence	Accurate Mass	Mass Defect	Isotopic Pattern Match	Consistent RT	Homologue (number; level)	MS^2 Fragments (number; type)	Library MS^2	Reference Standard
Level 1a	Confirmed by reference standard	✓	✓	✓	✓			✓	✓
Level 1b	Indistinguishable from reference standard	✓	✓	✓	✓			✓	✓
Level 2a	Probable by library spec. match	✓	✓	✓	✓			✓	
Level 2b	Probable by diagnostic fragmentation evidence	✓	✓	✓	✓	≥ 1; ≥ level 3	≥ 3; diagnostic		
Level 2c	Probable by diagnostic homologue evidence	✓	✓	✓	✓	≥ 2; ≥ level 2a	≥ 2; diagnostic		
Level 3a	Positional isomer candidates	✓	✓	✓	✓	≥ 1; ≥ level 3	≥ 1; subclass-aligned		
Level 3b	Fragmentation-based candidate	✓	✓	✓	✓	≥ 1; ≥ level 3	≥ 1; subclass-aligned		
Level 3c	Circumstantial candidate with fragmentation evidence	✓	✓	✓	✓	≥ 1; ≥ level 3	≥ 1; subclass-aligned (in silico)		
Level 3d	Circumstantial candidate with homologue evidence	✓	✓	✓	✓	≥ 2; ≥ level 2a			
Level 4	Unequivocal molecular formula	✓	✓	✓					
Level 5a	PFAS suspect screening exact mass match	✓ (suspect list match)							
Level 5b	Nontarget PFAS exact mass of interest	✓	✓			≥ 3	≥ 2; subclass-aligned (in silico)		

Figure 8: Identification confidence scale developed by Charbonnet et al. (2022)

3.3.1 Exploratory data analysis

After filtering and removing features labelled E (likely not PFAS), preliminary data visualisation was performed using principal component (PCA) and hierarchical cluster analysis (HCA).

Principal Component Analysis (PCA) is a statistical technique commonly used in data analysis to simplify and interpret complex datasets. PCA can be a powerful tool to uncover essential patterns and relationships hidden within experimental data. In essence, PCA reduces the dimensionality of the data while preserving the most critical information, allowing to visualize and understand data more effectively. By transforming the original variables into a new set of orthogonal components, called principal components, PCA highlights the main sources of variation in the dataset. This reduction in dimensionality helps identify underlying trends, clusters, or groupings in the data, aiding in the prioritisation of relevant features.

Similarly, **Hierarchical Cluster Analysis (HCA)** is used to identify patterns and similarities in large datasets. In HCA, samples are grouped into clusters based on similarities in their feature profiles (i.e., feature intensities), with the most similar samples forming a cluster and the least similar ones placed in separate clusters. The process involves creating a hierarchical tree-like structure, known as a dendrogram, which visually represents the clustering relationships. Dendrograms or heatmaps can be used to identify distinct groups or clusters within the data, helping uncover underlying structures and relationships between samples. The obtained information can help in the selection of relevant features for further investigation.

These were used to determine if clear differences in feature profiles across samples could be observed. More specifically, these techniques allow to determine if groups of samples can be formed based on their feature profiles, as well as to determine which features are causing the grouping. These will then be selected for a more detailed inspection. Principal Component Analysis (PCA) is a statistical technique commonly used in data analysis to simplify and interpret complex datasets. PCA can be a powerful tool to uncover essential patterns and relationships hidden within experimental data. In essence, PCA reduces the dimensionality of the data while preserving the most critical information, allowing scientists to visualize and comprehend their data more effectively. By transforming the original variables into a new set of orthogonal components, called principal components, PCA highlights the main sources of variation in the dataset. This reduction in dimensionality helps identify underlying trends, clusters, or groupings in the data, aiding in the identification of important biomarkers, chemical compounds, or experimental conditions. PCA is particularly valuable when dealing with large datasets with numerous variables, as it simplifies data interpretation and enables researchers to focus on the most influential factors impacting their scientific investigations.

In the case of results from C18 analysis, PCA showed that DW and GW samples were generally grouped together, except for S7.GW which clearly showed a different feature profile (see Figure 9). On the other hand, SW samples were more scattered, in particular S2.SWa and S4.SW which were clearly separated from the rest of the SW samples. Interestingly, DW from S4 appeared also to be separated from the rest of the DW samples, suggesting that this specific location has a different feature profile compared to other DW samples included in this study (see below for more information about this location). Sample S10.RBF, collected after riverbank infiltration, was also grouped with SW samples. Regarding features detected with MM, there was no clear distinction between sample types as in the case of C18 (see Figure 9). Only two drinking water samples, namely from locations S1 and S5 were clearly separated from the other samples. These findings seem to suggest that except for the above-mentioned DW samples, the differences in feature profiles between samples are less clearly defined with MM compared to C18.

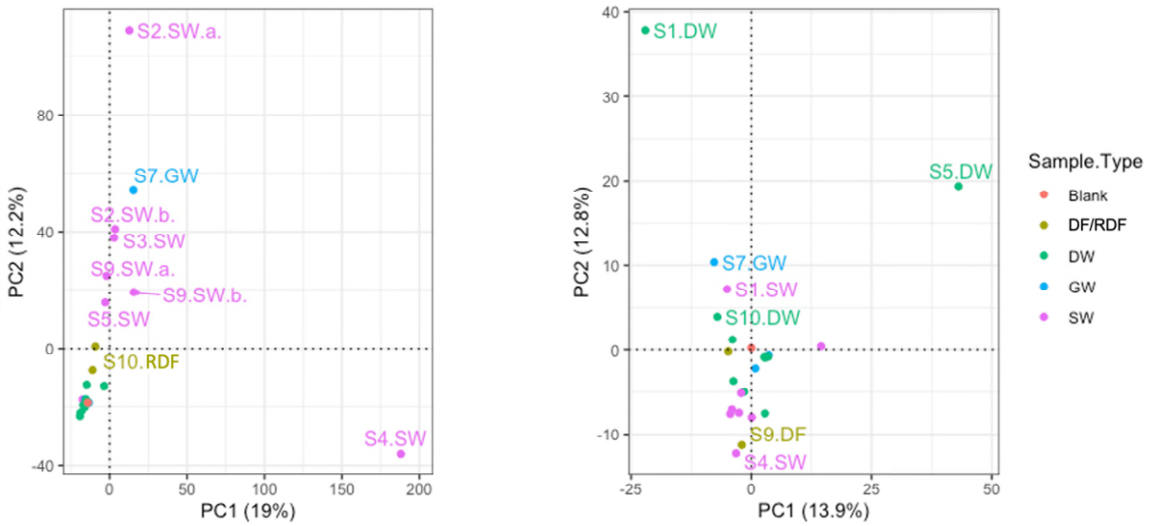


Figure 9: Principal component analysis (PCA) of samples analysed with C18 (left) and MM (right). Features scored E were removed prior to plotting. NB samples from location S11 were not analysed in SNTS.

For C18 analysis, hierarchical cluster analysis provided similar results to PCA, with most SW, as well as S7.GW, showing the largest differences (see Figure 10). Samples S10.RFB and S4.DW, which showed some separation in PCA (Figure 4), were here part of a larger cluster with various DW and GW samples. However, both were again separated from DW and GW samples at a higher level and showed a clearly distinct group of intense features absent in the other samples. Although less numerous compared to S4.DW, feature groups with higher intensity were present also in other DW samples. Interestingly, these features did not correspond to those that were most abundant in the corresponding samples of untreated SW or GW samples. These features are of course highly relevant as they are either not removed during treatment or consist of transformation products and will hence be investigated further in the following sections (see 3.3.2 for more details). Regarding features detected using MM (see Figure 11), as for PCA there was no clear distinction between sample types as is the case for C18. Interestingly, among sample showing the largest differences, there were various DW samples (i.e., S5.DW, S9.DW, S1.DW and S10.DW), all characterised by intense feature groups not present in other samples. This difference in clustering compared to C18 is likely due to the difference in chromatographic separation, as the MM column is expected to offer a better separation (i.e., retention) of more polar compounds. Furthermore, the mass range of features detected in MM is substantially smaller compared to C18 (i.e., 100-400 m/z *versus* 200-1200 m/z). The fact that a larger number of small and more polar features was detected with MM in DW samples might also be explained by the possible presence of transformation products, including shorter chain PFAS, formed during the drinking water treatment process (Li et al., 2020; Y. Q. Wang et al., 2022). Further investigation of features showing stable or increased intensities after drinking water treatment detected in both C18 and MM were further investigated in 3.3.2.

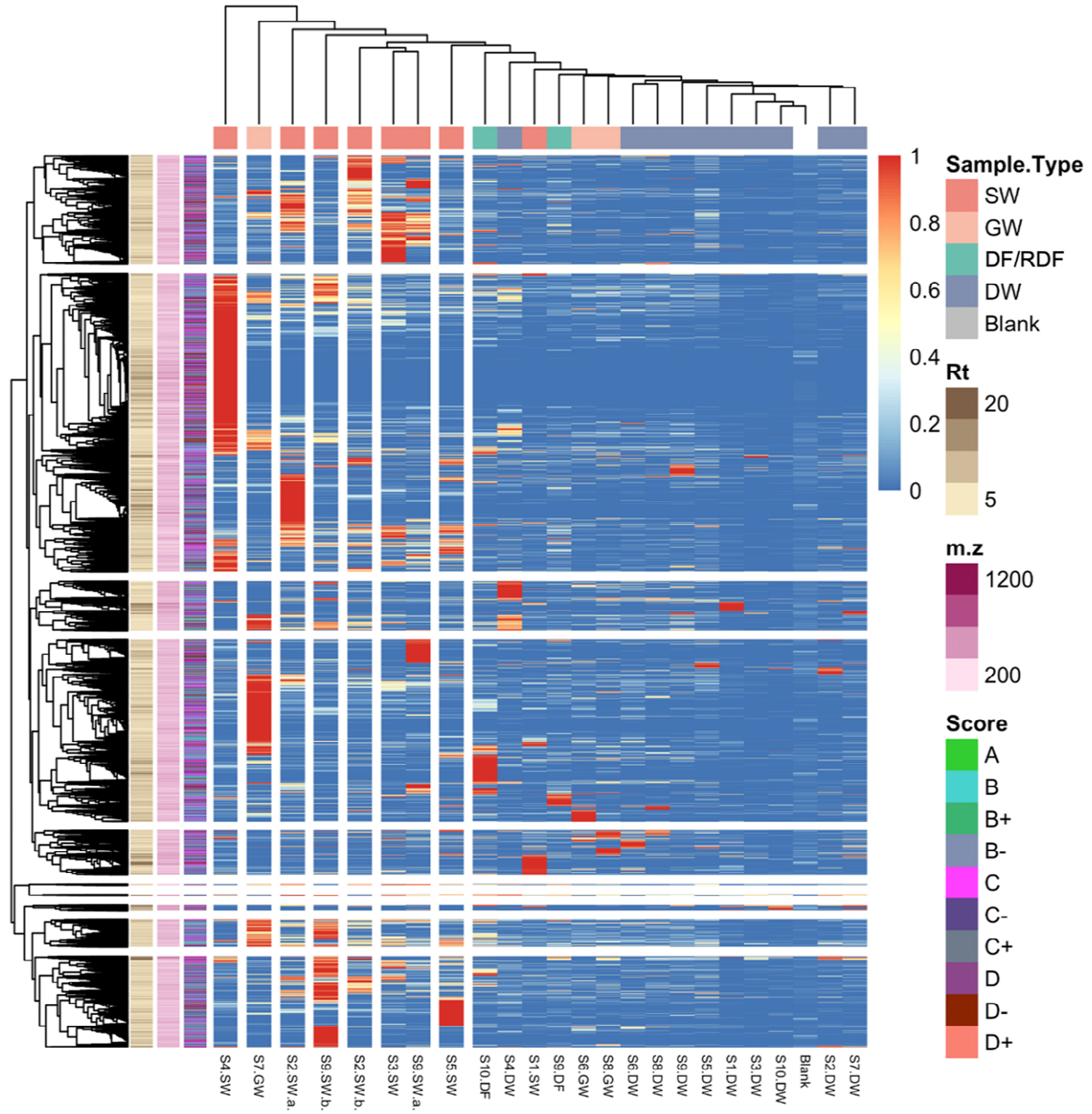


Figure 10: Heatmap of features detected with C18 (after filtering). Features scored E were removed prior to plotting. Retention times, measured m/z and scores provided by FluoroMatch are reported in the leftmost columns of the heatmap. NB samples from location S11 were not analysed in SNTS.

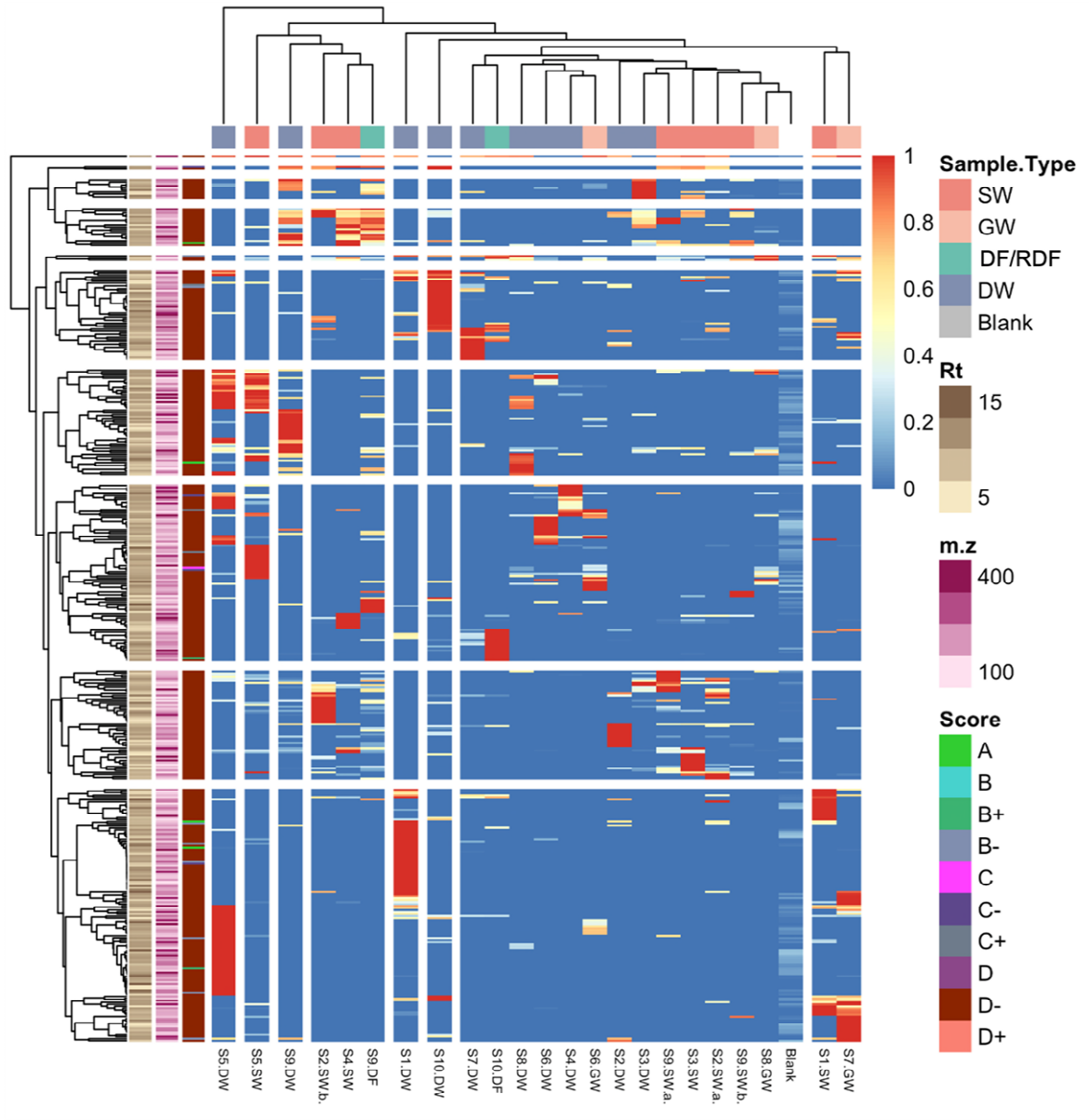


Figure 11: Heatmap of features detected with MM (after filtering). Features scored E were removed prior to plotting. Retention times, measured m/z and scores provided by FluoroMatch are reported in the leftmost columns of the heatmap. NB samples from location S11 were not analysed in SNTS.

3.3.1 Kendrick Mass Defect analysis

Kendrick Mass Defect analysis was performed to highlight the presence of homologous series (HS). The focus was set on CF_2, CF_2CH_2, OCF_2CFCF_3 and CF_2CFCl as repeating units. In the case of the latter two repeating units, no HS could be detected with C18. For CF_2 and CF_2CH_2 on the other hand, 26 and 36 potential HS were detected, respectively.

Kendrick Mass Defect

Kendrick Mass Defects (KMD) plots are commonly used to detect the presence of PFAS in SNTS analysis. This approach relies on the fact highly fluorinated chemicals tend to have a low or negative mass defect (i.e., the difference between the monoisotopic and nominal mass of a compound) due to the substitution of hydrogen atoms with fluorine. In addition, PFAS are often present in samples as homologous series (HS), which consist of chemicals of similar structure (e.g., same functional group such as a carboxylic acid) differing only in the length of the carbon chain. Commonly, HS of PFAS are represented by the number of CF_2 units which constitute the alkyl chain. These two characteristics can be exploited to prioritise features which are most likely PFAS. This is done by calculating KMD, which are nothing else than mass defects normalised by the mass of the repeating unit of interest (e.g., CF_2):

$$KMD = 414 - 413.97379 \times \frac{50}{49.9968} = 0.026$$

The example above illustrates how the KMD of PFOA is calculated based on the repeating unit CF_2. This is done for all features detected during the analysis and then KMDs are plotted against the nominal mass as shown in Figure 12. PFAS belonging to the same HS will have the same KMD (same value along the y-axis), yet will differ in their nominal mass due to different numbers of CF_2 units, which will translate in shifts along the x-axis of multiples of 50 (i.e., nominal mass of CF_2). However, features potentially belonging to the same HS also need to have coherent differences in retention time. In fact, when moving from right to left (i.e., increasing number of CF_2 units), the retention time of features also needs to increase.

Figure 12 illustrates the KMD plot obtained using CF_2 as a repeating unit. However, a more detailed analysis revealed that not all series were characterised by features whose retention time increased coherently with increasing nominal mass. Among those HS which did fulfil this condition, one consisted of legacy PFCA (number 55 in Figure 12) and one of legacy PFSA (number 31 in Figure 12), both of which were also detected during targeted analysis. Nevertheless, in the case of PFSA, features ranged from C2 to C8 and included three potentially relevant features not included in the targeted method, namely one with molecular formula $C_3HF_7O_3S$ (m/z of 248.9455 and retention time 3.656min), $C_6HF_{13}O_3S$ (m/z 398.9359 and retention time 6.871) and C8HF17O3S (m/z 498.9301 and retention time 11.14min), as shown in Table 2. The first feature was tentatively identified as heptafluoropropanesulfonic acid (PFPrS) (see Table 2). This feature was detected in all DW samples except for locations S6, S8. Interestingly, in location S3, PFPrS was detected only in DW samples but not in raw SW, suggesting that it might be formed during treatment. Nevertheless,

it cannot be excluded that this feature was not present in raw SW as transport times were not taken into account during sampling and only grab samples were collected. Furthermore, it cannot be excluded that the feature was not detected in SW due to potentially stronger signal suppression as a result of a more complex matrix compared to DW. The other two features were tentatively identified as branched isomers of PFHxS and PFOS, based on their fragmentation patterns and retention times which were in agreement with their linear analogues. In the case of the branched analogue of PFOS, this feature was detected only in samples from location S9 (both raw and treated drinking water). In the case of the HS attributed to PFCAs, an additional feature tentatively identified as a branched isomer of PFOA was detected (m/z 412.9658 and retention time 8.076 min). This feature was detected in most samples except for DW from locations S1 (SW), S6 (GW), S8 (GW) and S10 (riverbank filtration). Two more HS with coherent m/z and retention times were also highlighted. Namely, one with 7 features (of which 3 potential isomers) with coherent retention times and m/z ranging from 168.9887 to 368.9757, with confidence levels ranging from 3d to 5a (number 63 in Figure 12). However, this HS likely consisted of in-source fragmentation of potential PFAS features ranging from C3 to C7. The second consisted of 7 features (including 4 potential isomers) with m/z ranging from 167.0705 to 317.0655, all of which had confidence level 5a (number 299 in Figure 12). The remaining HS either had less than 3 features and/or were not characterised by consistent retention times. Regarding KMD analysis of data obtained with MM, no additional relevant features could be prioritised. In fact, no clear HS could be detected, likely due to the fact that only rather small and polar compounds were retained.

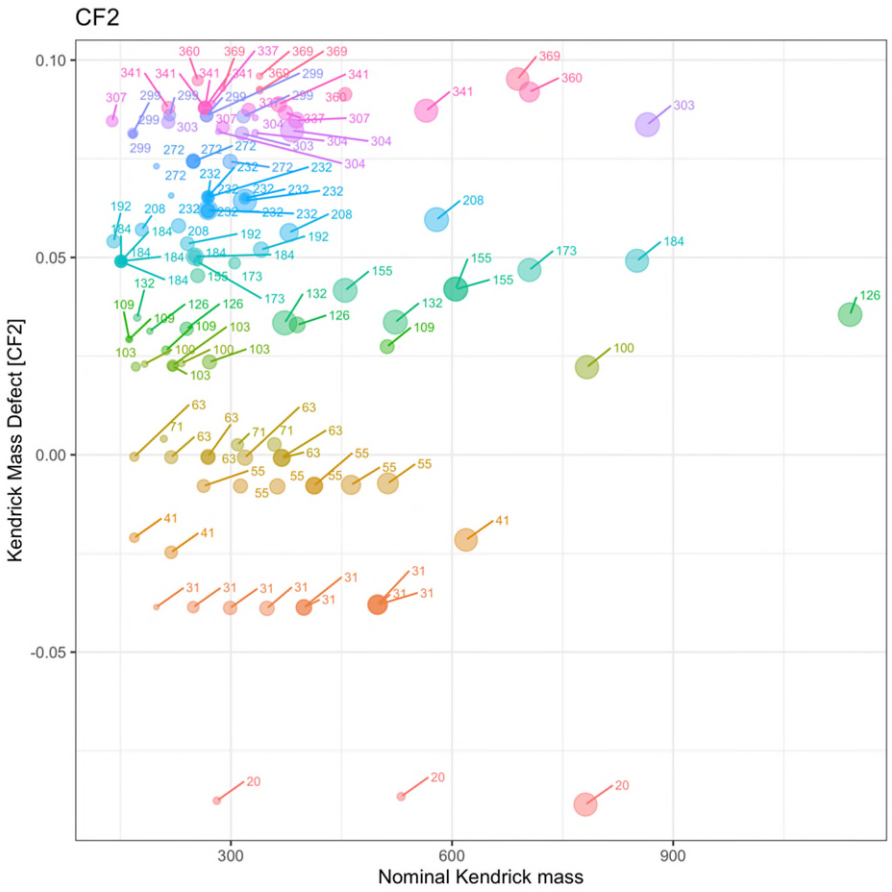


Figure 12: KMD plot of HS detected by FluoroMatch using CF2 as repeating unit following analysis with C18. Features with score E were not included here. Numbers indicate the corresponding HS number. NB samples from location S11 were not analysed in SNTS.

3.3.2 Raw versus treated drinking water feature profiles

Exploratory data analysis showed that, in particular using C18, clear differences in samples and feature profiles could be highlighted, while MM analysis highlighted a larger number of intense feature groups in drinking water samples. Following KMD analysis, homologous series of relevant features not detected with the targeted method were prioritised and tentatively identified. However, because KMD analysis focuses mainly on the presence of homologous series, it cannot be excluded that other PFAS (not part of a homologous series) might also be present in the samples. Consequently, a more detailed analysis was carried out by focusing on features showing either constant (i.e., little to no removal) or increasing (i.e., transformation products) intensity across treatment steps. Specifically, feature clusters showing constant or increasing intensities in HCA (see for an example), as well as features having a log2 fold-change larger or equal to 2 and a p-value < 0.05 in volcano plots were prioritised. Based on these two approaches, more than 6,000 features (labelled A to D) were prioritised, 312 of which were selected because of their significantly higher intensity in treated DW compared to raw GW or SW (i.e., volcano plots). Due to the large number of features still present, the focus was set on particularly intense features having a higher identification confidence (i.e., above label C) or for which an unequivocal molecular formula (containing fluorine) could be determined. Based on these criteria, 14 additional features were prioritised for further analysis and are reported in Table 2. Two features were

tentatively identified (level 2a) as N-methylperfluoropropane sulfonamido acetic acid (MeFPrSAA) and N-methylperfluorobutane sulfonamido acetic acid (MeFBSAA), which have both been reported as components of aqueous film forming foams (AFFFs) (Barzen-Hanson et al., 2017; Newton et al., 2017). The first feature was detected only in SW samples from locations S1 and S4, while the second was detected in a larger number of samples, including DW of locations S1 and S2 (see features 5 and 6 in Figure 14). An additional feature was tentatively identified with confidence level 2b, namely 6:2 Fluorotelomer sulfonamide betaine (6:2 FTAB, $C_{15}H_{19}F_{13}N_2O_4S$, feature 7 in Figure 14), which has been identified as a major component in AFFFs using PFOS alternatives (Moe et al., 2012; Shi et al., 2018). This feature was detected with increasing intensity (almost threefold) in the treatment train of location S5. While it is rather unlikely that this feature is being formed during treatment, considering that it is a precursor which would be expected to undergo oxidation to shorter/legacy PFAS, these results seem to suggest that it is not efficiently removed in location S5. An important characteristic about S5.SW is that it consists of a storage basin in the proximity of a PFAS manufacturing facility which has been brought into cause for PFAS contaminations (Gebbink & van Leeuwen, 2020). In addition to the latter location, this feature was also detected in SW from locations S2, S3 and S9, as well as GW from S7 and DF from S9. Five additional features (8 to 12 in Figure 14 and Figure 15) were tentatively identified with confidence levels ranging from 3b to 3d, as reported in Table 2. These consisted of one H-substituted PFSA ($C_4H_2F_8O_3S$), a suspected PFAS-ester ($C_6H_6F_6O2$), 2,2,2-Trifluoroethyl perfluorobutanesulfonate ($C_6H_2F_{12}O_3S$) and an unsaturated PFSA ($C_6HF_{11}O_3S$). Three additional features (13 to 15 in Figure 14) with identification level 4 (i.e., unequivocal molecular formula (Charbonnet et al., 2022)) were also prioritised, namely $C_6H_9F_3O_4S$ (tentatively annotated as n:2 $FTSO_2PA$, yet only based on an MS1 match), $C_9H_{10}F_2O_3S$ (tentatively annotated as 2,2-Difluoroethyl 4-methylbenzenesulfonate, yet only based on an MS1 match) and $C_3H_2F_4O_2$ (tentatively annotated as 2,3,3,3-tetrafluoropropanoic acid, based only on MS1 yet having a retention time of 8.74 which is consistent with TFA and PFBA in MM). Finally, four additional features (16 to 19 in Table 2 and Figure 15) were prioritised based on their detection frequency and high intensities in treated drinking water. However, no unequivocal molecular formula could be computed, and no tentative structure could be attributed.

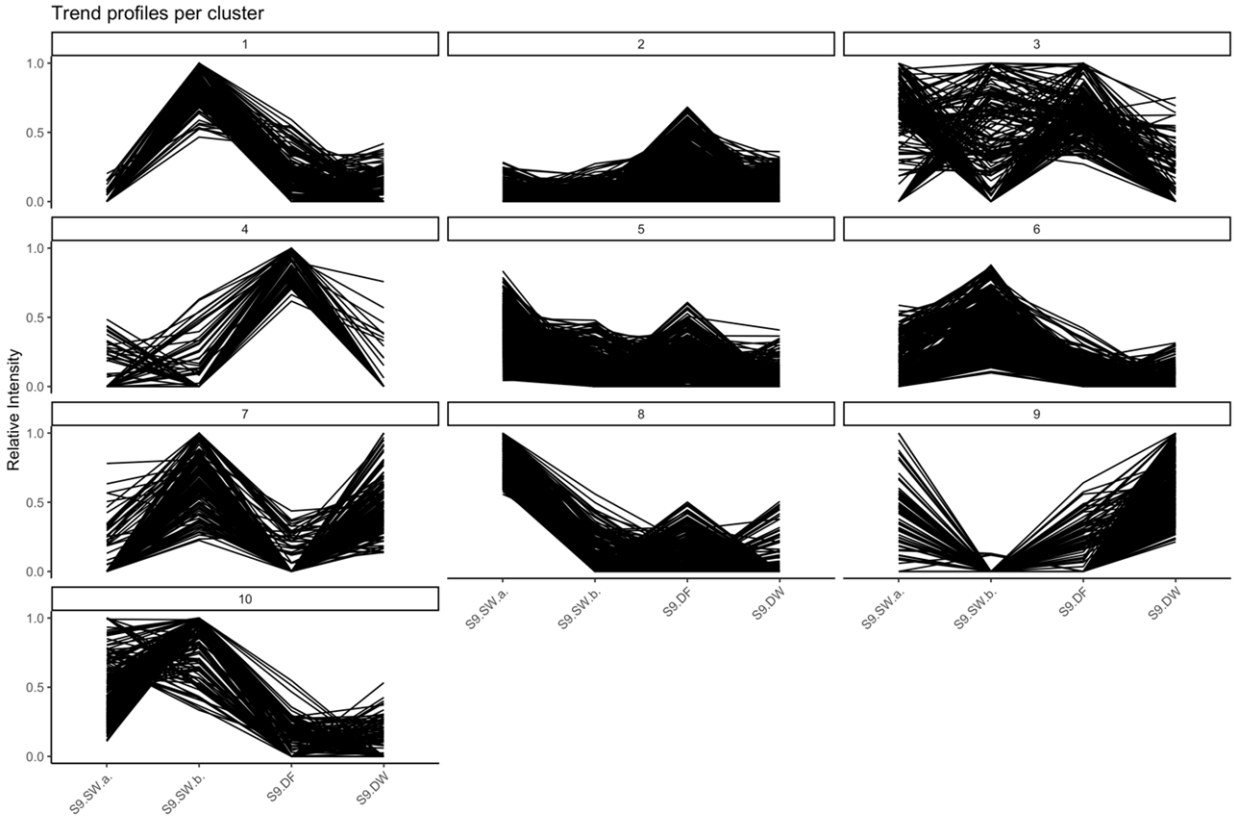


Figure 13: Trends in feature intensities across the various sampling points (i.e., treatment steps) of location S9.

Table 2: Overview of prioritised and tentatively identified features. NB samples from location S11 were not analysed in SNTS.

#	Method	Annotated structure	Formula	m/z	Rt	PFAS Class	Accurate Mass	Mass Defect	Isotopic Pattern Match	Consistent RT	HS (number; level)	Library MS2	Confidence Level
1	C18	Heptafluoropropanesulfonic acid (PFPrS)	$C_3HF_7O_3S$	248.9455	3.66	PFSA	✓	✓	✓	✓	7; 1a-2a	✓	2a
2	C18	Branched-Perfluorohexanesulfonic acid (Br-PFHxS)	$C_6HF_{13}O_3S$	398.9359	6.87	PFSA	✓	✓	✓	✓	7; 1a-2a	✓	2c
3	C18	Branched-Perfluorooctnesulfonic acid (Br-PFOS)	$C_8HF_{17}O_3S$	498.9301	11.14	PFSA	✓	✓	✓	✓	7; 1a-2a	✓	3d
4	C18	Branched-Perfluorooctanoic acid (Br-PFOA)	$C_8HF_{15}O_2$	412.9658	8.08	PFCA	✓	✓	✓	✓	6; 1a-2a	✓	2c
5	C18	N-methylperfluoropropane sulfonamido acetic acid (MeFPrSAA)	$C_6H_6F_7NO_4S$	319.9823	5.08	PFSM-carboxylic_acid	✓	✓	✓	✓	1; 2a	✓	2a
6	C18	N-methylperfluorobutane sulfonamido acetic acid (MeFBSAA)	$C_7H_6F_9NO_4S$	369.9791	6.04	PFSM-carboxylic_acid	✓	✓	✓	✓	1; 2a	✓	2a
7	C18	6:2 Fluorotelomer sulfonamide betaine	$C_{15}H_{19}F_{13}N_2O_4S$	569.0779	10.66		✓	✓	✓	✓			2b
8	C18	1,1,2,2,3,3,4,4-Octafluorobutane-1-sulphonic acid	$C_4H_2F_8O_3S$	280.9516	3.73	(PFSA-H)	✓	✓	✓	-	-		3b
9	MM	Methyl 2,3,3,4,4,4-hexafluoro-2-methylbutanoate	$C_6H_6F_6O_2$	223.0196	2.13	PFAS-ester	✓	✓	✓				3b
10	C18	2,2,2-Trifluoroethyl perfluorobutanesulfonate	$C_6H_2F_{12}O_3S$	380.9451	5.08		✓	✓	✓				3c
11	C18		$C_6HF_{11}O_3S$	360.9390	6.03	PFSA-unsaturated				✓			3c
12	C18	8:2 Fluorotelomer sulfonamide betaine	$C_{17}H_{19}F_{17}N_2O_4S$	669.0724	17.53		✓	✓	✓				3d
13	C18	FT-Sulfone (n:2 FTSO2PA)	$C_6H_9F_3O_4S$	233.0083	1.77	Fluorotelomer	✓	✓	✓	-	-		4
14	C18	2,2-Difluoroethyl 4-methylbenzenesulfonate	$C_9H_{10}F_2O_3S$	235.0240	4.69		✓	✓	✓	-	-		4
15	MM	2,3,3,3-tetrafluoropropanoic acid	$C_3H_2F_4O_2$	144.9916	8.74		✓	✓	✓	✓			4
16	C18		$C_3HF_7O_5S$	280.9352	4.39		✓	✓	✓	-			5a
17	C18		$C_6H_8F_4O_2$	233.0448	1.60		✓	✓	✓	-			5a
18	C18		$C_5H_4F_9N$	361.1039	1.92		✓	✓	✓	-	-		5a
19	C18		$C_3HF_7O_5S$	280.9352	4.39		✓	✓	✓				5a

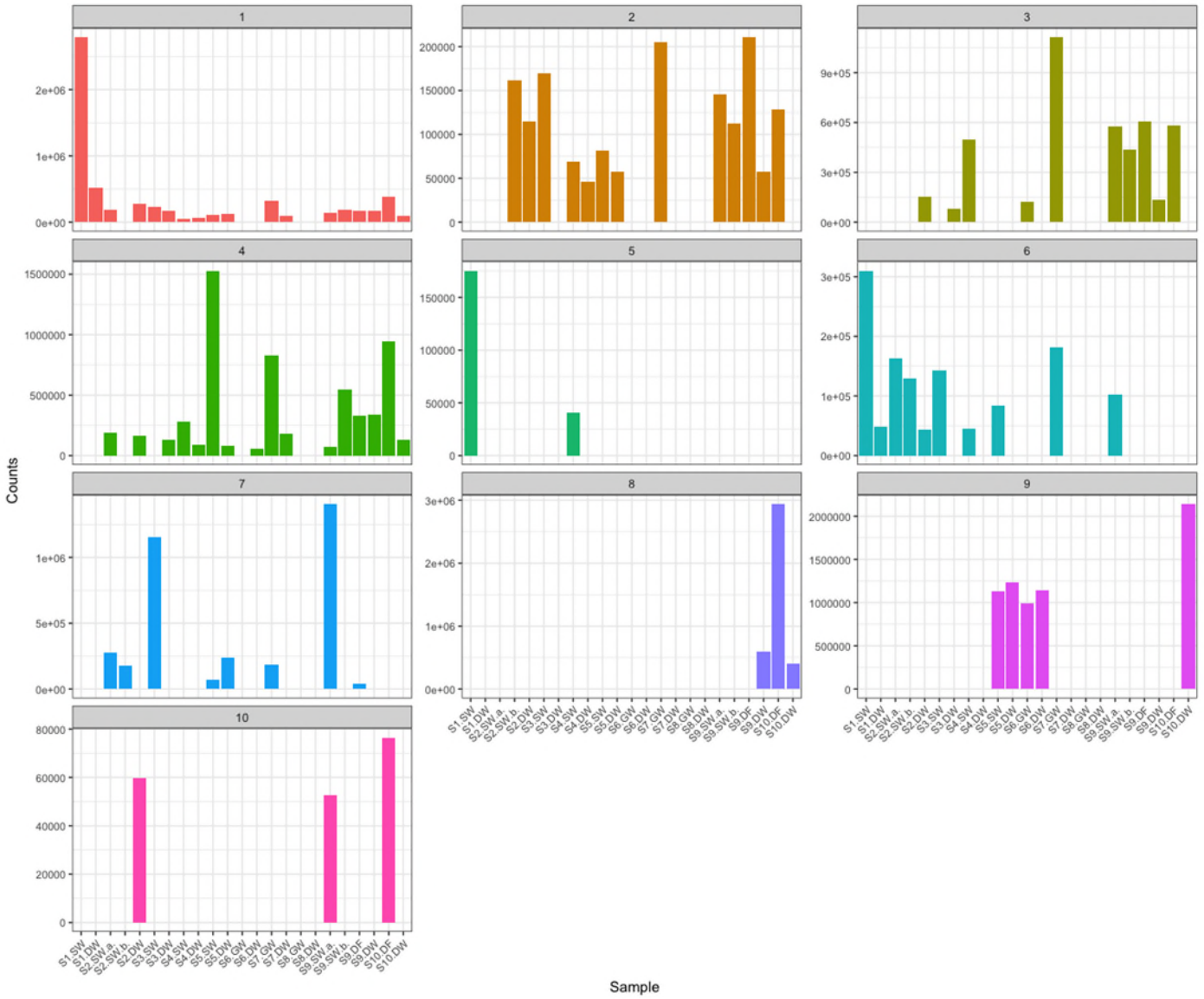


Figure 14: Occurrence of features prioritised due to their constant or increasing intensity during drinking water treatment. NB samples from location S11 were not analysed in SNTS.

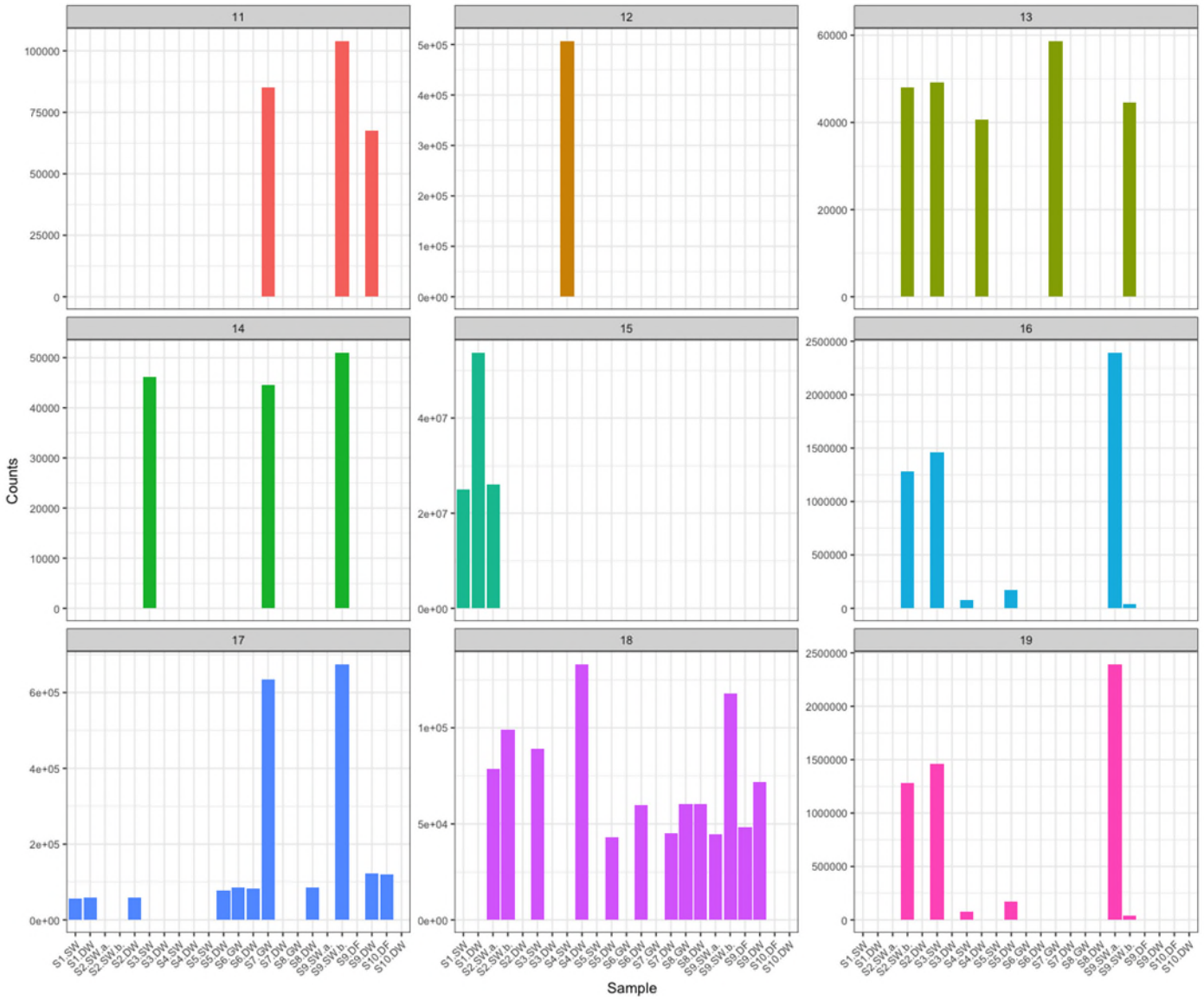


Figure 15: Occurrence of features prioritised due to their constant or increasing intensity during drinking water treatment. NB samples from location S11 were not analysed in SNTS.

3.4 Comparison between targeted and suspect and non-target screening

As expected, concentrations of PFAS measured with the targeted method, as well as the number of potential PFAS features detected with SNTS, were higher in surface water compared to groundwater samples, except for groundwater location S7 which showed both higher legacy PFAS concentration and a conspicuous number of features which could be potentially PFAS (although not all had high identification confidence levels). Regarding the specific case of drinking water samples, a large number of potential PFAS features was still present in samples from location S9 (S9-DW), although most had low identification confidence levels. This is however in line with findings from targeted analysis, which also highlighted higher PFAS concentrations at this location. With respect to sample location S10, which involves riverbank infiltration, overall high legacy PFAS concentrations were observed in the raw (source) water but the number of potential PFAS features still found in DW was substantially lower, when compared to S9. Overall, it can be concluded that SNTS highlighted the presence of various potentially relevant features which were not covered by the targeted method. Furthermore, these features were not all effectively removed through the considered treatment processes. Whilst some of these could be tentatively identified with a rather high confidence levels, others still require additional analysis to both increase the confidence that they are actually PFAS and eventually confirm their identity.

Meer informatie
Dr Frederic Béen
T +31 30 606 9748
E frederic.been@kwrwater.nl

PO Box 1072
3430 BB Nieuwegein
The Netherlands

4 Conclusions

This project covered both the development and application of comprehensive analytical methods to monitor PFAS in raw and treated drinking water. Method development involved the validation of two targeted methods for the analysis of ultra-short and short/long-chain PFAS, respectively. Furthermore, these two methods were also implemented in full scan mode to acquire suspect and non-target screening (SNTS) data, with the goal of detecting additional compounds which were not included in the targeted methods. Targeted analyses showed that 10 out of the 36 compounds monitored could be detected with high frequency in most samples, including drinking water. However, in drinking water the measured concentrations were often substantially lower compared to untreated water, except for TFA. Regarding the 4 PFAS included in the EFSA guidelines, concentrations in DW samples were consistently below the 4.4 ngPEQ/L threshold except for locations S5 and S9 which showed slightly higher concentrations.

Among the ultra-short chain PFAS, TFA could be detected (semi-quantitatively) at high levels in almost all samples and was only limitedly removed and, in some cases, concentrations in drinking water were higher than those in raw water. However, as highlighted earlier, these results were obtained from the analysis of individual samples and retention times were not accounted for. A more thorough assessment of removal and/or potential formation of PFAS should include a larger number of samples as well as a sampling scheme which accounts for residence times. A selected number of PFAS precursors was also included in the targeted analysis, yet only 6:2 FTS and FBSA were detected in a few samples.

These results might suggest that the presence of PFAS precursors in drinking water sources is limited, however the compounds included in the targeted method represent only a fraction of all the precursors potentially present in the aquatic environment. To determine if additional PFAS, including precursors, were presented in raw and finished drinking water, SNTS analyses were carried out using FluoroMatch for data processing. Most PFCA and PFSA measured in the targeted method were also detected with high confidence using the implemented workflow, providing some validation with respect to its capability to detect known PFAS. Through the use of KMD-plots and various data analysis strategies, additional features could be prioritised and tentatively identified, including 7 with relatively high confidence (level 2a-2c), some of which were detected also in drinking water. Whilst among these there were branched isomers of PFHxS, PFOA and PFOS, which are already included in current monitoring programs, other features such as MeFPrSAA, MeFBSAA and 6:2 Fluorotelomer sulfonamide betaine were tentatively identified. Because of their structure, these compounds fall under the category of PFAS precursors and, not surprisingly, were detected exclusively in samples from surface water.

These results indicate that other precursors, beyond the conventional ones included in targeted methods (including the one used here), might be present in surface waters and could contribute to measured legacy PFAS concentrations. Further analyses should be carried out to confirm their identity, including that of other prioritised

features, and subsequently, a decision should be taken on whether to include them in future targeted monitoring programs.

The results obtained in this study have shown that the combination of targeted and SNTS provides valuable information about the occurrence of PFAS in drinking water sources. While targeted methods are essential to quantify low levels of PFAS, they cannot provide a comprehensive picture of all compounds potentially present. On the other hand, SNTS can help expand the number of PFAS that can be monitored, in particular thanks to the growing number of processing tools and workflows which facilitate data analysis. Nevertheless, the complexity of the data together with the high number of detected features, make this a time consuming exercise. Yet, by screening samples more regularly and by focusing only on those features showing consistent patterns (e.g., constant or increasing intensity) over time (i.e., multiple sampling campaigns or longitudinal studies), it would be possible to narrow down even further the number of features to be investigated further. Whilst more samples would need to be analysed, in the end, there would be a gain in processing time because prioritisations would be based on repeated measurements over time rather than incidental analyses as it happens currently.

5 References

Ateia, M., Maroli, A., Tharayil, N., & Karanfil, T. (2019). The overlooked short- and ultrashort-chain poly- and perfluorinated substances: A review. *Chemosphere*, *220*, 866–882. https://doi.org/10.1016/j.chemosphere.2018.12.186

Barzen-Hanson, K. A., Roberts, S. C., Choyke, S., Oetjen, K., McAlees, A., Riddell, N., McCrindle, R., Ferguson, P. L., Higgins, C. P., & Field, J. A. (2017). Discovery of 40 Classes of Per- and Polyfluoroalkyl Substances in Historical Aqueous Film-Forming Foams (AFFFs) and AFFF-Impacted Groundwater. *Environmental Science & Technology*, *51*(4), 2047–2057. https://doi.org/10.1021/acs.est.6b05843

Bugsel, B., & Zwiener, C. (2020). LC-MS screening of poly- and perfluoroalkyl substances in contaminated soil by Kendrick mass analysis. *Analytical and Bioanalytical Chemistry*, *412*(20), 4797–4805. https://doi.org/10.1007/s00216-019-02358-0

Charbonnet, J. A., McDonough, C. A., Xiao, F., Schwichtenberg, T., Cao, D., Kaserzon, S., Thomas, K. V., Dewapriya, P., Place, B. J., Schymanski, E. L., Field, J. A., Helbling, D. E., & Higgins, C. P. (2022). Communicating Confidence of Per- and Polyfluoroalkyl Substance Identification via High-Resolution Mass Spectrometry. *Environmental Science & Technology Letters*, *9*(6), 473–481. https://doi.org/10.1021/acs.estlett.2c00206

Proposal for a Directive of the European Parliament and of the Council on the quality of water intended for human consumption, 6060/1/20 REV 1 (2020). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CONSIL:ST_6060_20 20_REV_1&from=EN

Eschauzier, C., Raat, K. J., Stuyfzand, P. J., & De Voogt, P. (2013). Perfluorinated alkylated acids in groundwater and drinking water: Identification, origin and mobility. *Science of The Total Environment*, *458–460*, 477–485. https://doi.org/10.1016/j.scitotenv.2013.04.066

Gebbink, W. A., & van Leeuwen, S. P. J. (2020). Environmental contamination and human exposure to PFASs near a fluorochemical production plant: Review of historic and current PFOA and GenX contamination in the Netherlands. *Environment International*, *137*, 105583. https://doi.org/10.1016/j.envint.2020.105583

Glüge, J., Scheringer, M., Cousins, I. T., DeWitt, J. C., Goldenman, G., Herzke, D., Lohmann, R., Ng, C. A., Trier, X., & Wang, Z. (2020). An overview of the uses of per- and polyfluoroalkyl substances (PFAS). *Environmental Science: Processes & Impacts*. https://doi.org/10.1039/D0EM00291G

Hu, X. C., Andrews, D. Q., Lindstrom, A. B., Bruton, T. A., Schaider, L. A., Grandjean, P., Lohmann, R., Carignan, C. C., Blum, A., Balan, S. A., Higgins, C. P., & Sunderland, E. M. (2016). Detection of Poly- and Perfluoroalkyl Substances (PFASs) in U.S. Drinking Water Linked to Industrial Sites, Military Fire Training Areas, and Wastewater Treatment Plants. *Environmental Science & Technology Letters*, *3*(10), 344–350. https://doi.org/10.1021/acs.estlett.6b00260

Jacob, P., Barzen-Hanson, K. A., & Helbling, D. E. (2021). Target and Nontarget Analysis of Per- and Polyfluoralkyl Substances in Wastewater from Electronics Fabrication Facilities. *Environmental Science & Technology*, *55*(4), 2346–2356. https://doi.org/10.1021/acs.est.0c06690

Jeong, Y., Da Silva, K. M., Iturrospe, E., Fuiji, Y., Boogaerts, T., van Nuijs, A. L. N., Koelmel, J., & Covaci, A. (2022). Occurrence and contamination profile of legacy and emerging per- and polyfluoroalkyl substances (PFAS) in Belgian wastewater using target, suspect and non-target screening approaches. *Journal of Hazardous Materials*, *437*, 129378. https://doi.org/10.1016/j.jhazmat.2022.129378

Kaboré, H. A., Vo Duy, S., Munoz, G., Méité, L., Desrosiers, M., Liu, J., Sory, T. K., & Sauvé, S. (2018). Worldwide drinking water occurrence and levels of newly-identified perfluoroalkyl and polyfluoroalkyl substances. *Science of The Total Environment*, *616–617*, 1089–1100. https://doi.org/10.1016/j.scitotenv.2017.10.210

Koelmel, J. P., Paige, M. K., Aristizabal-Henao, J. J., Robey, N. M., Nason, S. L., Stelben, P. J., Li, Y., Kroeger, N. M., Napolitano, M. P., Savvaides, T., Vasiliou, V., Rostkowski, P., Garrett, T. J., Lin, E., Deigl, C., Jobst, K., Townsend, T. G., Godri Pollitt, K. J., & Bowden, J. A. (2020). Toward Comprehensive Per- and Polyfluoroalkyl Substances Annotation Using FluoroMatch Software and Intelligent High-Resolution Tandem Mass Spectrometry Acquisition. *Analytical Chemistry*, *92*(16), 11186–11194. https://doi.org/10.1021/acs.analchem.0c01591

Koelmel, J. P., Stelben, P., McDonough, C. A., Dukes, D. A., Aristizabal-Henao, J. J., Nason, S. L., Li, Y., Sternberg, S., Lin, E., Beckmann, M., Williams, A. J., Draper, J., Finch, J. P., Munk, J. K., Deigl, C., Rennie, E. E., Bowden, J. A., & Godri Pollitt, K. J. (2022). FluoroMatch 2.0—Making automated and comprehensive non-targeted PFAS annotation a reality. *Analytical and Bioanalytical Chemistry*, *414*(3), 1201–1215. https://doi.org/10.1007/s00216-021-03392-7

Li, F., Duan, J., Tian, S., Ji, H., Zhu, Y., Wei, Z., & Zhao, D. (2020). Short-chain per- and polyfluoroalkyl substances in aquatic systems: Occurrence, impacts and treatment. *Chemical Engineering Journal*, *380*, 122506. https://doi.org/10.1016/j.cej.2019.122506

Moe, M. K., Huber, S., Svenson, J., Hagenaars, A., Pabon, M., Trümper, M., Berger, U., Knapen, D., & Herzke, D. (2012). The structure of the fire fighting foam surfactant Forafac®1157 and its biological and photolytic transformation products. *Chemosphere*, *89*(7), 869–875. https://doi.org/10.1016/j.chemosphere.2012.05.012

Nakayama, S. F., Yoshikane, M., Onoda, Y., Nishihama, Y., Iwai-Shimada, M., Takagi, M., Kobayashi, Y., & Isobe, T. (2019). Worldwide trends in tracing poly- and perfluoroalkyl substances (PFAS) in the environment. *TrAC Trends in Analytical Chemistry*, *121*, 115410. https://doi.org/10.1016/j.trac.2019.02.011

Newton, S., McMahen, R., Stoeckel, J. A., Chislock, M., Lindstrom, A., & Strynar, M. (2017). Novel Polyfluorinated Compounds Identified Using High Resolution Mass Spectrometry Downstream of Manufacturing Facilities near Decatur, Alabama. *Environmental Science & Technology*, *51*(3), 1544–1552. https://doi.org/10.1021/acs.est.6b05330

NORMAN Network. (2023). *NORMAN Suspect List Exchange*. https://www.norman-network.com/?q=suspect-list-exchange

OECD. (2023). *OECD Portal on Per and Poly Fluorinated Chemicals—Comprehensive Global Database of PFASs*. https://www.oecd.org/chemicalsafety/portal-perfluorinated-chemicals/

RStudio Team. (2021). *RStudio: Integrated Development Environment for R* [Computer software]. http://www.rstudio.com/

Sadia, M., Nollen, I., Helmus, R., ter Laak, T. L., Béen, F., Praetorius, A., & van Wezel, A. P. (2023). Occurrence, Fate, and Related Health Risks of PFAS in Raw and Produced Drinking Water. *Environmental Science & Technology*. https://doi.org/10.1021/acs.est.2c06015

Scheurer, M., Nödler, K., Freeling, F., Janda, J., Happel, O., Riegel, M., Müller, U., Storck, F. R., Fleig, M., Lange, F. T., Brunsch, A., & Brauch, H.-J. (2017). Small, mobile, persistent: Trifluoroacetate in the water cycle – Overlooked sources, pathways, and consequences for drinking water supply. *Water Research*, *126*, 460–471. https://doi.org/10.1016/j.watres.2017.09.045

Schrenk, D., Bignami, M., Bodin, L., Chipman, J. K., Mazo, J. del, Grasl-Kraupp, B., Hogstrand, C., Hoogenboom, L. (Ron), Leblanc, J.-C., Nebbia, C. S., Nielsen, E., Ntzani, E., Petersen, A., Sand, S., Vleminckx, C., Wallace, H., Barregård, L., Ceccatelli, S., Cravedi, J.-P., … Schwerdtle, T. (2020). Risk to human health related to the presence of perfluoroalkyl substances in food. *EFSA Journal*, *18*(9), e06223. https://doi.org/10.2903/j.efsa.2020.6223

Schulze, T., Meier, R., Alygizakis, N., Schymanski, E., Bach, E., LI, D. H., lauperbe, raalizadeh, Tanaka, S., & Witting, M. (2021). *MassBank/MassBank-data: Release version 2021.12* [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.5775684

Schymanski, E. L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H. P., & Hollender, J. (2014). Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environmental Science & Technology*, *48*(4), 2097–2098. https://doi.org/10.1021/es5002105

Schymanski, E., Zhang, J., Thiessen, P., Chirsir, P., Kondic, T., & Bolton, E. (2023). *Per- and polyfluoroalkyl substances (PFAS) in PubChem: 7 million and growing*. ChemRxiv. https://doi.org/10.26434/chemrxiv-2023-j823z

Shi, G., Xie, Y., Guo, Y., & Dai, J. (2018). 6:2 fluorotelomer sulfonamide alkylbetaine (6:2 FTAB), a novel perfluorooctane sulfonate alternative, induced developmental toxicity in zebrafish embryos. *Aquatic Toxicology*, *195*, 24–32. https://doi.org/10.1016/j.aquatox.2017.12.002

Drinking Water Directive, 2020/2184 (2020). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32020L2184&from=EN

United Nations Environment Programme. (2019). *Stockholm Convention*. http://www.pops.int/

van der Aa, M., & te Biesebeek, J. D. (2021). *Analyse bijdrage drinkwater en voedsel aan blootstelling EFSA-4 PFAS in Nederland en advies drinkwaterrichtwaarde*. RIVM.

Wang, Y. Q., Hu, L.-X., Liu, T., Zhao, J.-H., Yang, Y.-Y., Liu, Y.-S., & Ying, G.-G. (2022). Per- and polyfluoralkyl substances (PFAS) in drinking water system: Target and non-target screening and removal assessment. *Environment International*, *163*, 107219. https://doi.org/10.1016/j.envint.2022.107219

Wang, Z., Cousins, I. T., Scheringer, M., & Hungerbühler, K. (2013). Fluorinated alternatives to long-chain perfluoroalkyl carboxylic acids (PFCAs), perfluoroalkane sulfonic acids (PFSAs) and their potential precursors. *Environment International*, *60*, 242–248. https://doi.org/10.1016/j.envint.2013.08.021

Wang, Z., Schymanski, E., & Williams, A. (2018). *S25 | OECDPFAS | List of PFAS from the OECD* [dataset]. https://doi.org/10.5281/zenodo.3653165

I Appendix

I.I Methods for liquid chromatography Optimization

An initial LC method was based on a wide range of literature [7], [9], [16], [29]–[31], [33], [41], [46], [50], [51]. The analytical column chosen was the XBridge® BEH C18 to achieve a reversed phase separation and column temperature was kept at 25 °C. Injection volume was 10 µL and flow rate 0.3 mL/min. Initial eluent composition was A: 5 mM NH4Ac in UP-water and B: 5 mM NH4Ac in methanol. For both solutions buffer concentrations of 2 mM and 1 mM NH4Ac were tested. Then the gradient was optimized starting with a linear increase from 5% B to 100% B in 10 minutes and kept at 100% for 4 minutes. Afterwards 80:20 MeOH:ACN %(v/v) was tested as solvent for eluent B as well. The optimal parameters were chosen by highest signal and lowest baseline noise level of PFAS in a 1 µg/L solution of PFAC30PAR in 20:80 MeOH:UP-water. A gradient elution method was developed based on the separation of 30 analytes in the previously mentioned solution and a maximum analysis time of 25 minutes. Finally, the methanol percentage in the standard was tested at 25, 50 and 75% as well.

I.II Methods for mass spectrometry optimization

Orbitrap Fusion parameters were then optimized to achieve an optimal MS detection sensitivity for all 36 PFAS analytes. The following parameters were optimized. Sheath gas, varied from 20 to 60 Arb in steps of 10 Arb; Auxiliary gas, from 0 to 20 Arb in steps of 5 Arb; Sweep gas, 2, 5 and 10 Arb; Ion transfer tube and vaporizer temperature, from 200 to 350 °C in steps of 50 °C and finally RF-Lens, 30 to 80 % in steps of 10%. The optimal settings were chosen from the average highest sensitivity of the 30 PFAS Analytes combined. Other parameters were set without further optimization. Resolution was set at 120,000, the automatic gain control at 200,000, injection time at 100 ms, negative voltage at 2,500 V and scan range at 150-1000 m/z. These parameters were not optimized further.

I.III C18 method validation results for drinking water

Compound	LOD	LOQ	Reporting limit	Recovery [5 ng/L]	
	[ng/L]	[ng/L]	[ng/L]	[%]	RSD [%]
PFBA	0.15	0.45	0.50	102	4
PFPeA	0.07	0.21	0.50	101	5
PFBS	0.04	0.12	0.50	106	1
4:2-FTSA	0.06	0.18	0.50	104	2
PFHxA	0.06	0.19	0.50	88	5
HFPO-DA	0.10	0.29	0.50	103	6
PFPeS	0.04	0.13	0.50	101	3
FBSA	0.03	0.10	0.50	60	6
PFHpA	0.06	0.18	0.50	102	1
NaDONA	0.07	0.20	0.50	126	5

PFHxS	0.04	0.11	0.50	104	2
6:2-FTSA	0.14	0.42	0.50	103	2
PFOA	0.04	0.13	0.50	104	1
PFHpS	0.04	0.11	0.50	104	2
FHxSA	0.02	0.06	0.50	66	2
PFNA	0.04	0.11	0.50	103	1
PFOS	0.04	0.12	0.50	103	2
9Cl-PF3ONS	0.04	0.13	0.50	93	2
8:2-FTSA	0.13	0.38	0.50	99	2
PFDA	0.03	0.10	0.50	100	1
PFNS	0.02	0.07	0.50	86	2
N-MeFOSAA	0.09	0.27	0.50	100	3
FOSA	0.03	0.10	0.50	108	2
N-EtFOSAA	0.15	0.44	0.50	102	2
PFUnDA	0.09	0.27	0.50	102	2
PFDS	0.12	0.36	0.50	95	7
11Cl-PF3OUdS	0.11	0.34	0.50	57	10
PFDoDA	0.11	0.33	0.50	100	3
PFTrDA	0.73	2.20	3.0	48	43
PFTeDA	0.84	2.52	3.0	96	34

I.IV C18 method validation results for surface water

Compound	**LOD**	**LOQ**	**Reporting limit**	**Recovery [5 ng/L]**	
	[ng/L]	**[ng/L]**	**[ng/L]**	**[%]**	**RSD [%]**
PFBA	0.18	0.53	0.50	117	4
PFPeA	0.14	0.42	0.50	103	2
PFBS	0.14	0.43	0.50	100	2
4:2-FTSA	0.06	0.19	0.50	99	4
PFHxA	0.13	0.39	0.50	105	2
HFPO-DA	0.08	0.23	0.50	99	2
PFPeS	0.07	0.22	0.50	112	5
FBSA	0.06	0.19	0.50	88	5
PFHpA	0.05	0.16	0.50	101	1
NaDONA	0.04	0.11	0.50	104	2
PFHxS	0.05	0.16	0.50	102	1
6:2-FTSA	0.06	0.17	0.50	102	1
PFOA	0.13	0.38	0.50	102	1
PFHpS	0.04	0.12	0.50	73	4
FHxSA	0.07	0.22	0.50	96	4
PFNA	0.04	0.13	0.50	103	1
PFOS	0.03	0.10	0.50	103	1
9Cl-PF3ONS	0.05	0.14	0.50	130	5
8:2-FTSA	0.09	0.26	0.50	104	1

PFDA	0.03	0.09	0.50	101	1
PFNS	0.05	0.15	0.50	73	7
N-MeFOSAA	0.11	0.33	0.50	104	2
FOSA	0.04	0.13	0.50	103	1
N-EtFOSAA	0.19	0.56	0.50	99	5
PFUnDA	0.04	0.13	0.50	103	2
PFDS	0.07	0.20	0.50	93	8
11Cl-PF3OUdS	0.06	0.19	0.50	74	10
PFDoDA	0.06	0.18	0.50	102	2
PFTrDA	0.39	1.16	3.0	116	42
PFTeDA	0.37	1.11	3.0	101	7

I.V MM method validation results for drinking water

Compound	LOD	LOQ	Reporting Limit	Recovery [5 ng/L]	
	[ng/L]	[ng/L]	[ng/L]	[%]	RSD [%]
PFMeSAm	0.13	0.38	5.0	15	10
PFEtSAm	0.11	0.34	1.0	82	7
TFA	1.79	5.36	20	57	2
F3-MSA	0.07	0.22	2.0	54	12
PFPrA	0.34	1.02	10	102	5
PFEtSA	0.10	0.31	1.0	73	4

I.VI MM method validation results for surface water

Compound	LOD	LOQ	Reporting limit	Recovery [5 ng/L]	
	[ng/L]	[ng/L]	[ng/L]	[%]	RSD [%]
PFMeSAm	0.13	0.38	5.0	14	11
PFEtSAm	0.08	0.24	1.0	81	8
TFA	1.79	5.36	20	103	7
F3-MSA	0.51	1.53	2.0	172	5
PFPrA	0.68	2.05	10	210	3
PFEtSA	0.09	0.27	1.0	74	9

I.VII Results of the liquid chromatography optimization

Starting from the initial LC method described in chapter 3.4, first the gradient was optimized. The goal was to achieve as much separation between the 30 PFAS analytes in the 1 µg/L standard solution within 25 minutes. The first combined extracted ion chromatogram obtained is shown in Figure 4A, this is from the non-optimized initial method. Elution of the first compound was at 5 minutes, while the rest elutes between 8 and 12 minutes. This signifies that a

high percentage of organic modifier is necessary to elute most PFAS in the standard. Additionally, all analytes were detected with sufficient sensitivity. Requiring no major method overhauls for the mass spectrometer. The optimal gradient starts with 30 %B for 1 minute, then ramps up to 60 %B in 2 minutes and stays level for 4 minutes, then ramps up to 90 %B non-linearly in 9 minutes, increases linearly to 100 %B in 1 minute and lastly stays at 100 %B for 3 minutes. The chromatogram obtained from this method is presented in Figure 4B. On the left a combined chromatogram shows the separation of all compounds. Although some analytes coelute, these can still be identified and separated through the high mass accuracy of the MS. The chromatogram is split up into five colours that represent five classes of PFAS. Blue is carboxylic acids, red is sulfonic acids, green is sulfonamides, purple is novel PFAS and orange is fluorotelomers. From these a clear increase in retention time with increasing chain length can be observed. This observation was expected to decreasing polarity with increasing chain length.

After the gradient was optimized, the effect of buffer concentration in the eluent on analyte signal was tested. To increase the analyte signal and reduce baseline noise as much as possible. The average peak area of all analytes was compared to determine the optimal parameters. Table 3 shows those values. An ammonium acetate buffer concentration in the eluent of 2 mM provided the largest average intensity. Additionally, the baseline noise with 2 mM was lower than 5 mM, however, higher than for 1 mM. Although the difference in average intensity between 1 mM and 2 mM was minimal the difference for individual compounds was significant with intensity differences higher than 30%, which is shown in Attachment 2. 12 compounds showed a minimal increase in intensity from 1% to 24%, summing up to 182%. While the other 18 compounds showed a large decrease in intensity from -2% up to -78%, summing up to -618%. Because 1 mM showed a significant decrease in signal intensity compared to 2 mM, the 2 mM buffer concentration was chosen as optimal.

Table 3: Average intensity for each buffer concentration

Buffer concentration	Average Intensity
5 mM	292588
2 mM	396292
1 mM	385076

Table 4: Average peak area for each methanol percentage in sample composition

Methanol percentage [%]	Average peak area
25	757020
50	823010
75	836860

Next the chromatographic separation was further optimized by addition of acetonitril to eluent B. Adding 20% acetonitril increased the separation between analytes. This is depicted in Figure 4C and was compared to 4B. Most overlapping peaks were consequently baseline separated and signal of the more apolar and higher mass compounds was increased. Therefore 80:20 MeOH:ACN %(v/v) with 2 mM NH4Ac was chosen as optimal eluent B.

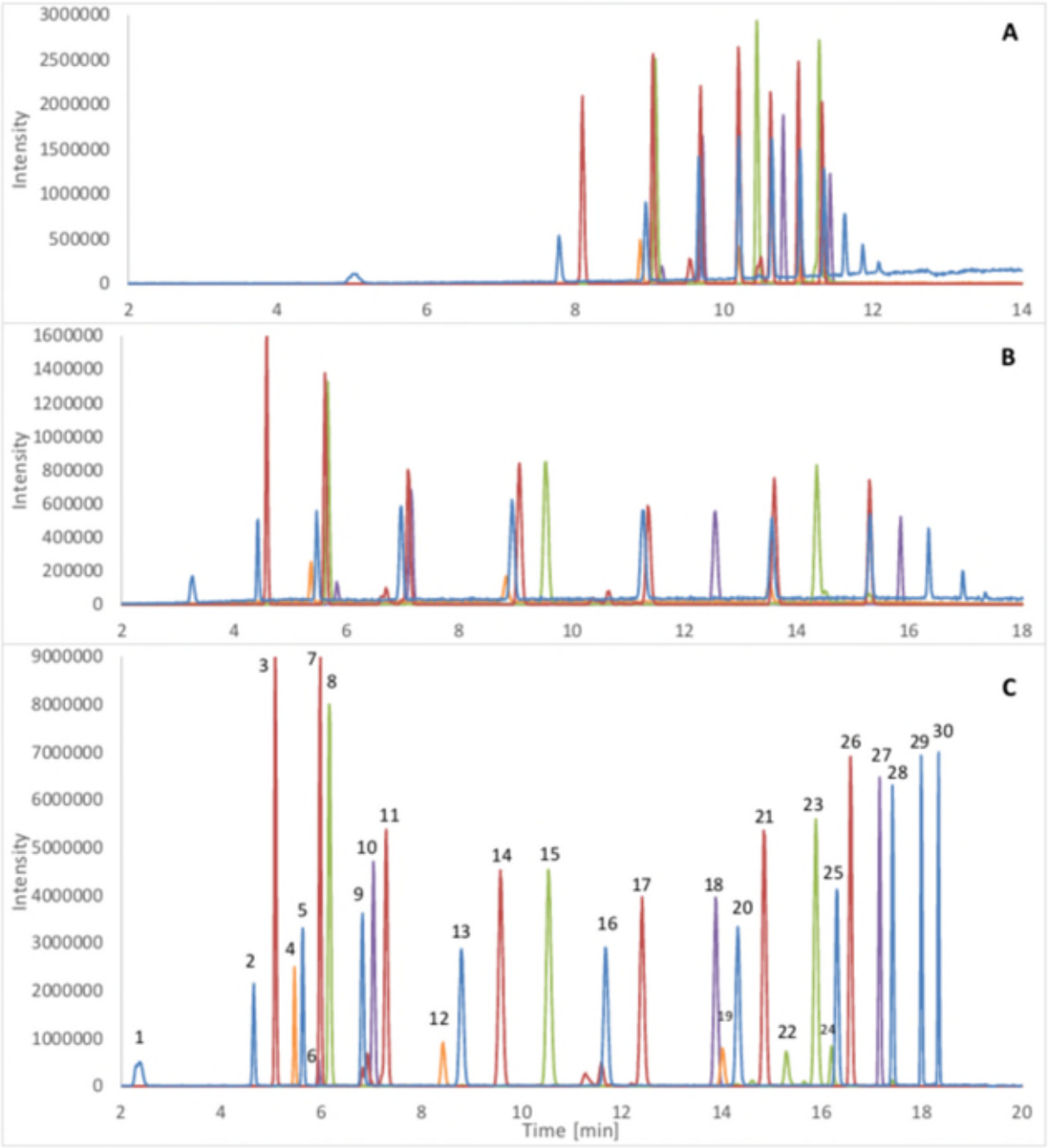


Figure 16: Chromatograms indicating three stages of LC optimization. A: Chromatogram acquired with initial method. B: Chromatogram after gradient and buffer concentration optimization. C: Chromatogram of the final method. The colours indicate the PFAS-classes. Blue is carboxylic acids, red is sulfonic acids, green is sulfonamides, purple is novel PFAS and orange is fluorotelomers. 1: PFBA, 2: PFPeA, 3: PFBS, 4: 4:2-FTS, 5: PFHxA, 6: HFPO-DA, 7: PFPeS, 8: FBSA, 9: PFHpA, 10: NaDONA, 11: PFHxS, 12: 6:2-FTS, 13: PFOA, 14: PFHpS, 15: FHxSA, 16: PFNA, 17: PFOS, 18: 9Cl-PF3ONS, 19: 8:2-FTS, 20: PFDA, 21: PFNS, 22: N-MeFOSAA, 23: FOSA, 24: N-EtFOSAA, 25: PFUnDA, 26: PFDS, 27: 11Cl-PF3OUdS, 28: PFDoDA, 29: PFTrDA and 30: PFTeDA

Then the percentage of methanol in the sample was optimized. Three standards in 25%, 50% and 75% methanol respectively were analysed. 75% methanol showed the highest average peak area as presented in Table 3. The full data for each individual analyte is depicted in Attachment 3. When comparing the individual difference per analyte the difference in peak area was negligible for most compounds except for PFBA, PFDoDA, PFTrDA and PFTeDA. For

PFBA and PFTeDA, 50% methanol was more beneficial. While 75% is more beneficial for the other two analytes. At 25% methanol the long chain PFAS from chain lengths of C10 and higher showed a big decrease in peak area. However, the same sample will be used for the ultra-short chain PFAS method. That method will require a relatively low methanol percentage in the sample. Therefore, as a compromise, 50% methanol was chosen as sample composition.

To separate and measure ultrashort-chain PFAS a different separation mechanism than reversed phase was required. This method was created with the Trinity Acclaim P1 analytical column, which is a trimodal mixed mode column with reversed phase, cation- and anion exchange functionality. The eluent used for this method is A: 2 mM NH4Ac in UP-water and B: 40 mM NH4Ac in 80:20 ACN:UP-water %(v/v). Only the gradient was adjusted to achieve sufficient separation. The MS settings were kept the same as from the short-long chain PFAS method, except for the RF-Lens which was set to 30%. This was done because a lower percentage showed a higher signal for compounds with a lower molecular weight. The optimal elution gradient for the chosen eluent and column was as follows. 15 to 90 %B in 7 minutes, kept at 90 %B for 3 minutes, ramped up to 100 %B in 1 minute and kept at 100 %B for 3 minutes. A chromatogram is shown in Figure 5. PFMeSAm, PFEtSAm, TFA and F3-MSA are fully separated, only PFPrA and PFEtS show a slight overlap. TFA was also found in high concentrations in a blank solution. This is likely due to its common appearance in various water sources and air deposition, as well as being a compound present in the tune mix for the mass spectrometer [7], [52], [53]. Due to this abundance of the analyte the LOD will be much higher than for other PFAS. After this method was optimized, method validation was attempted. However, data varied largely between duplicates of samples. To the best of our knowledge this may result from matrix effects, caused by co-elution of inorganic ions, which are often observed in mixed-mode chromatography [54]. However, due to time constraints the actual cause could not be investigated. Therefore, the method was left at this stage during this research.

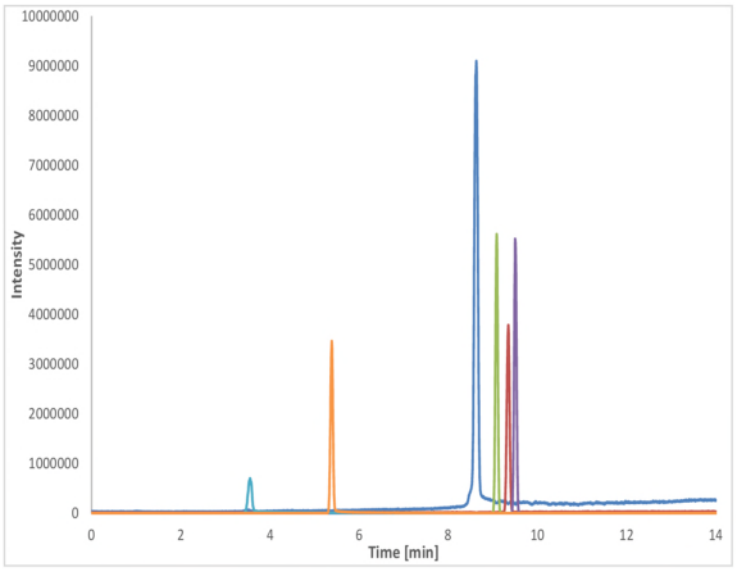


Figure 17: Ultrashort-chain PFAS chromatogram. The cyan trace is PFMeSAm, orange is PFEtSAm, blue is TFA, green is F3-MSA, Red is PFPrA and purple is PFEtS.

I.VIII Results of the mass spectrometry optimization

Six MS parameters were optimized. Namely sheath gas, auxiliary gas and sweep gas expressed in arbitrary units (arb); ion transfer tube and vaporizer temperature; and the RF-lens expressed in percentages (%). The average peak areas and a graphical overview of each parameter value separated by analyte are shown in Attachment 3. The average peak area combined for each analyte separated by MS-parameter is depicted in Figure 17. For all parameters except the RF-lens the optimal value is clear through comparison of the average peak area. 30 Arb for sheath gas, 5 Arb for auxiliary gas, 5 Arb for sweep gas, 200 °C for the ion transfer tube temperature and 300 °C for the vaporizer temperature. Although 20 Arb for sheath gas provided the highest peak area, this low setting may not fully evaporate the mobile phase in the ionization source, therefore, 30 Arb was chosen. The RF-lens shows the highest average peak area at 60%. However, when looking at the effect on individual compounds, smaller carboxylic acids and HFPO-DA are more affected by this parameter than larger compounds and sulfonic acids. A selection was made to show this graphically in Figure 19. Short chain carboxylic acids like PFBA and PFPeA, and HFPO-DA show the largest peak are at low RF-lens percentages. This quickly decreases with increasing RF-lens percentage. Long chain carboxylic acids like PFOA show a shift towards higher percentages as optimum. Sulfonic acids and other PFAS provide a distribution like the one shown for PFDS. Although on average 60% was optimal, because PFBA, PFPeA and HFPO-DA showed a very large dependence on the RF-lens setting, while other compounds do not, an RF-lens value of 50% was chosen as optimal value for this method. The resulting chromatogram from all optimized settings is shown in Figure 16C together with the first chromatogram recorded. Showcasing the result of the finalized method to measure short to long chain PFAS. Thereby achieving the first part of this research to create a method to separate as many PFAS, from short- to long chain, as possible.

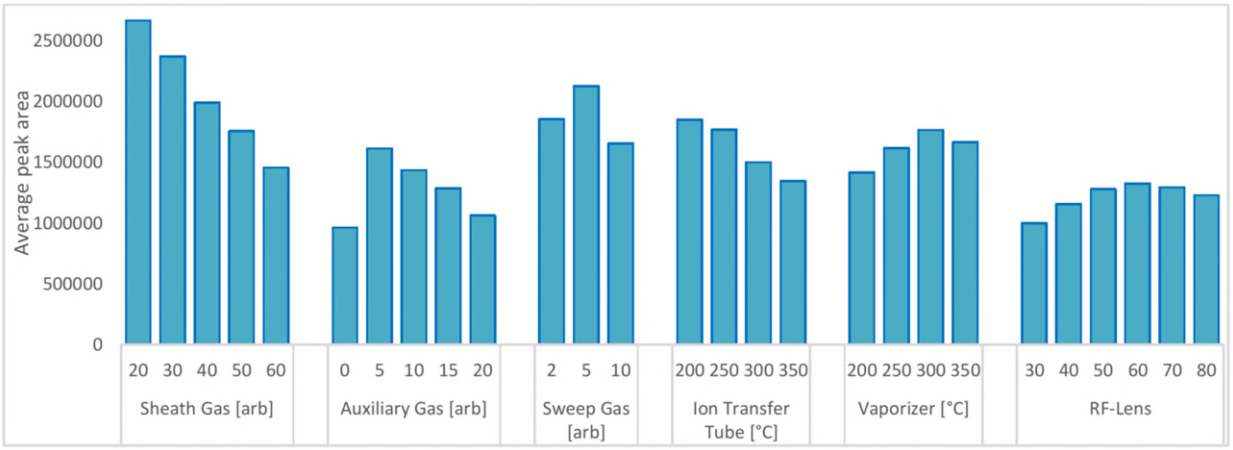


Figure 18: Bar graph depicting total average peak areas for each MS-parameter setting.

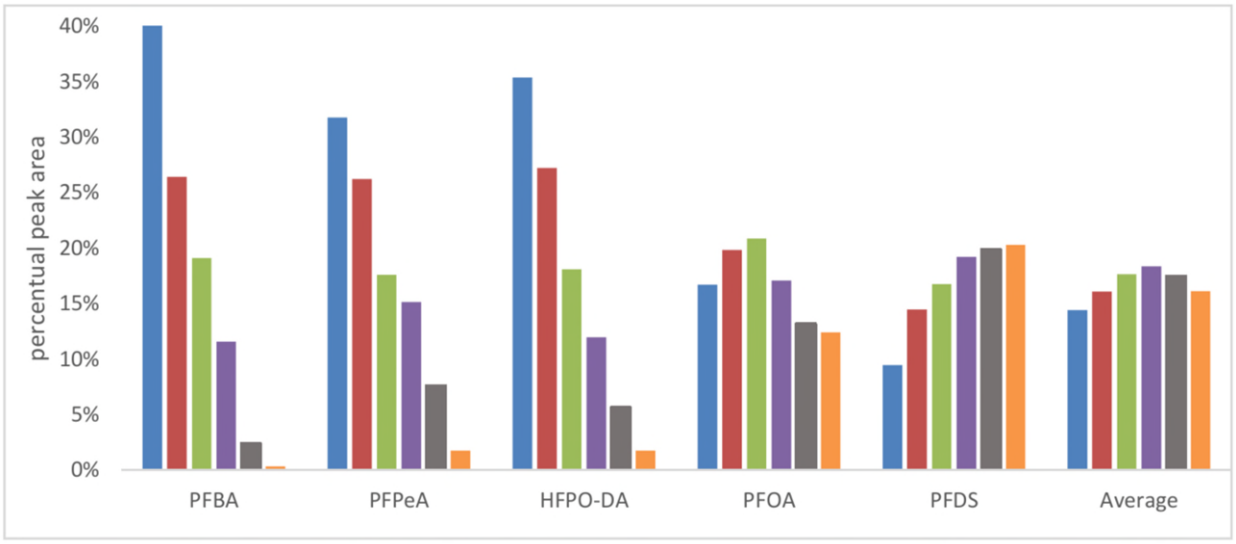


Figure 19: Graphical representation of the percentual peak areas (as fraction of the sum peak area of all values for that analyte) of PFBA, PFPeA, HFPO-DA, PFOA, PFDS and the total average for RF-Lens values of 30% (Blue), 40% (Red), 50% (Green), 60% (Purple), 70% (Grey) and 80% (Orange) respectively.

To conclude the full LC-MS optimization the final method is summarized in Table 4. This includes all optimized and pre-set settings. A chromatogram of the short-long chain PFAS method is provided in Figure 4C. The chromatogram for the ultrashort chain PFAS method is shown in Figure 17.