

BTO 2018.029 | Mei 2018

BTO rapport

Selecteren van relevante
nieuwe stoffen in
drinkwater op basis van
non-target-screening

BTO

Selecteren van relevante nieuwe stoffen in drinkwater op basis van non-target-screening

BTO 2018.029 | mei 2018

Opdrachtnummer

401399/001

Projectmanager

Stefan Kools

Opdrachtgever

Vitens - Speerpuntonderzoek

Kwaliteitsborger(s)

Annemarie van Wezel

Auteur(s)

Thomas ter Laak,
Andrea Brunner,
Rosa Sjerps,
Dennis Vughs,
Mario Maessen (Vitens),
Milou Dingemans,
Bernard Bajema (Vitens)

Verzonden aan

Bernard Bajema, Ronny Bosch, Mario Maessen en
Martin de Jonge (Vitens).

Een jaar na publicatie is deze rapportage openbaar.

Jaar van publicatie

2018

Meer informatie

Thomas ter Laak
T 030 60 657
E Thomas.ter.laak@kwrwater.nl

Keywords

non target screening, prioriteren en
identificeren van onbekende
stoffen, grondwater,
transformatieproducten

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

KWR

Watercycle
Research
Institute

BTO 2018.029 | maart 2018 © KWR

Alle rechten voorbehouden.

Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of enig andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

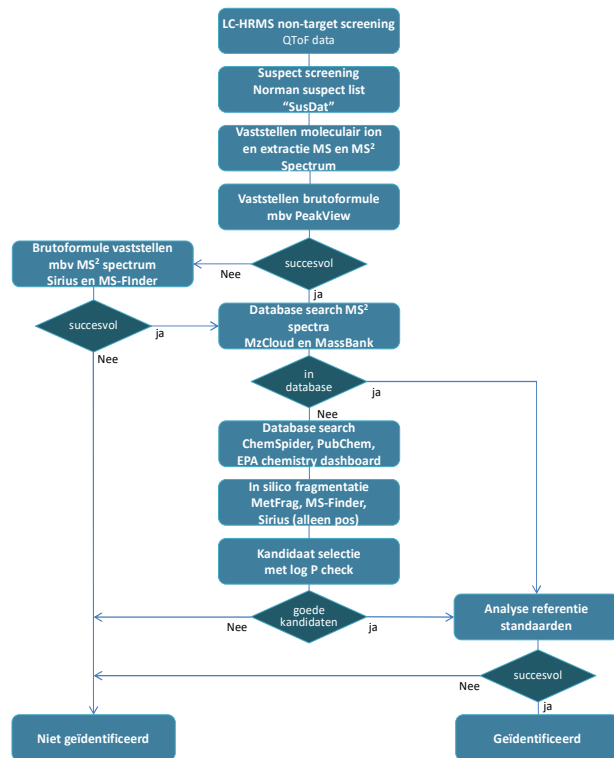
BTO Managementsamenvatting

Selecteren van relevante nieuwe stoffen in drinkwater op basis van non-target-screening

Auteur(s)

Thomas ter Laak, Andrea Brunner, Rosa Sjerps, Dennis Vughs, Mario Maessen, Milou Dingemans en Bernard Bajema

Bij non-target chemische screening van kwetsbare grondwaterwinningen wordt een scala aan onbekende stoffen aangetroffen. Het is niet mogelijk alle stoffen te identificeren en toxicologisch te beoordelen. In deze studie is een prioriteringssysteem ontworpen om de meest relevante stoffen te selecteren. De prioritering is gebaseerd op de frequentie en intensiteit van het aantreffen en de (mogelijke) antropogene herkomst. Daarnaast is een workflow ontwikkeld om te bepalen wat nodig en effectief is om de geprioriteerde stoffen te kunnen identificeren. De prioritering en identificatie-workflow helpt bij het bepalen welke stoffen het meest relevant zijn en het structureren van vervolgonderzoek naar de identiteit van grote hoeveelheden onbekende stoffen in (bronnen van) drinkwater. Daarmee biedt de systematiek handvatten voor het inrichten en/of updaten van een risico getuurd monitoringsprogramma.



Workflow voor identificatie van geprioriteerde onbekende stoffen

Belang: hoe bepaal je welke onbekende stoffen het meest relevant zijn?

Non-target screening methoden toegepast in (bronnen) van drinkwater leiden tot de detectie van een breed scala aan onbekende organische verbindingen. Het identificeren en kwantificeren van deze stoffen is zeer arbeidsintensief. Het is daarom van belang om de meest relevante stoffen te selecteren alvorens verder onderzoek naar de identiteit van deze stoffen wordt gedaan.

Aanpak: Prioriteren en identificeren van onbekende stoffen.

Het voorliggend onderzoek tracht onbekende stoffen te prioriteren en identificeren op basis van een database van non-target analyses uitgevoerd op 394 monsters in 29 winningen van waterbedrijf Vitens. De waargenomen onbekende stoffen zijn geprioriteerd op basis van het voorkomen in ruw en rein water (frequentie en de intensiteit), de mogelijk antropogene herkomst, hun relatie met het landgebruik boven de betreffende winning en of deze stoffen worden toegevoegd of gevormd tijdens de zuivering. Vervolgens is een workflow ontwikkeld om onbekende stoffen efficiënt te kunnen identificeren. In deze worden diverse databases geraadpleegd en worden modellen toegepast om fragmentatie van stoffen te voorspellen en te vergelijken met meetgegevens. Bovendien worden kandidaat stoffen beknopt toxicologisch gescreend om te bepalen welke onbekende stoffen in aanmerking komen voor verdere identificatie. Vervolgens worden monsters geselecteerd voor her-analyse en standaarden voor aanvullende analyses voor respectievelijk betere identificatie of bevestiging (of afwijzing) van de identiteit van de kandidaat stof.

Resultaten: 22 geprioriteerde stoffen nader onderzocht

De prioritering van ruim 11.000 onbekende stoffen leidde tot de selectie van 12 onbekende stoffen in de positieve ionisatie modus en 10 stoffen in de negatieve ionisatie modus. Voor 12 van deze 22 geprioriteerde stoffen zijn kandidaat stoffen geselecteerd. Deze zijn geanalyseerd ter bevestiging van de identiteit. Geen van de kandidaten bleek daadwerkelijk de aangetroffen stof te zijn. Vier van de onbekende stoffen zijn nader bestudeerd door geselecteerde monsters opnieuw te analyseren met analytische apparatuur met hogere resolutie. Hoewel dit niet direct tot identificatie leidde kon de brutoformule beter worden bepaald en kandidaten beter worden geselecteerd. De identificatie van de geprioriteerde stoffen is lastig. Dit komt vermoedelijk vooral doordat aangetroffen stoffen transformatieproducten zijn die niet of beperkt aanwezig zijn in databases en waarvan standaarden ontbreken. Deze evaluatie toont het belang aan van internationale databases aan voor de selectie van geschikte kandidaat stoffen in grondwater.

Implementatie: onderzoek naar onbekende stoffen in bronnen, identificatie en toxicologische evaluatie.

Non-target analyses stellen ons in staat een breed scala aan stoffen in het (water) milieu te meten. De identificatie van deze stoffen is – zoals ook uit deze studie blijkt - ingewikkeld. In dit onderzoek is een prioriteringsmethodiek ontwikkeld om gestructureerd de meest relevante onbekende stoffen te selecteren voor nader onderzoek en eventuele identificatie. Deze methodiek biedt handvatten om onbekende stoffen op basis van voorkomen en antropogene kenmerken te prioriteren. Deze stoffen kunnen vervolgens geselecteerd worden voor toxicologische evaluatie en worden opgenomen in de monitoring. Daarmee kunnen deze methoden worden ingezet om te bepalen welke stoffen op welke locaties relevant kunnen zijn en om risico gestuurde monitoring – programma's vorm te geven. De non-target screening is ook op andere wijze in te zetten, zoals voor het bepalen van meer generieke eigenschappen van een monster (herkomst, waterkwaliteit, zuiveringsrendement, correlaties met resultaten van bio-assays). Daarmee bieden de ontwikkelde methoden ook handvatten voor het invullen van een risico gestuurd monitoring programma's die in de revisie van Annex II van het Drinkwaterbesluit wordt beschreven.

Inhoud

Inhoud	2
1 Inleiding	4
1.1 Achtergrond	4
1.2 Doel en onderzoeksvragen	4
1.3 Aanpak	5
2 Duiding onbekende stoffen	6
2.1 Experimentele opzet en achtergrond	6
2.2 Evaluatie herkomst stoffen ruwwater	8
2.3 Correlatie met hydrologische kenmerken en landgebruik van de winningen	9
2.4 Correlatie met macroparameters van het gewonnen water	11
2.5 Correlatie bekende en onbekende stoffen	12
2.6 Vergelijk voorkomen in kwetsbare en niet kwetsbare winningen	15
2.7 Chemische samenstelling van onbekende stoffen	16
2.8 Verwijderingsefficiëntie	16
2.9 Transformatieproducten	17
2.10 Van stofeigenschap naar duiding	20
3 Prioriteren van onbekende stoffen	21
3.1 Aanpak	21
3.2 Data voorbereiding	22
3.3 Prioriteren van onbekende stoffen	23
3.4 Prioriteren van Toxische eigenschappen	26
3.5 Totaalscore prioritering	30
3.6 Geprioriteerde features	31
4 Identificatie en evaluatie geprioriteerde stoffen	33
4.1 Workflow identificatie geprioriteerde stoffen	33
4.2 Toepassing	34
4.3 Resultaten identificatie geprioriteerde stoffen	35
4.4 Bevestiging geprioriteerde suspects	38
4.5 Mogelijkheden voor verdere identificatie	40
4.6 Analyse van geprioriteerde stoffen met de Orbitrap Fusion massaspectrometer	42
4.7 Herziende workflow voor identificatie van onbekende verbindingen	46
4.8 Toxicologische evaluatie van geprioriteerde stoffen	48
5 Conclusies en aanbevelingen	50
5.1 Conclusies	50
5.2 Aanbevelingen	51

6	Literatuur	53
7	Bijlagen	55
	Bijlage I Correlaties tussen CMR classificaties en op bruto formule gebaseerde toxiciteitsklassen van CMR geclassificeerde stoffen uit ECHA.	56
	Bijlage II Overzicht referentie-standaarden voor identificatie van kandidaat stoffen	57
	Bijlage III Mogelijke transformatie producten geprioriteerde stoffen	58
	Bijlage IV Analyse methode non-target screening Orbitrap Fusion	60
	Monstervoorbewerking vaste fase extractie	60
	Monstervoorbewerking directe injectie	60
	Orbitrap Fusion non-target screening methode	60
	Bijlage V Chromatogrammen en massaspectra Orbitrap Fusion analyse	61
	Resultaten Orbitrap Fusion analyse feature 2431 (P10)	61

1 Inleiding

1.1 Achtergrond

Via verschillende routes komen antropogene stoffen in het milieu terecht. Screening naar bekende stoffen (doelstoffen) geeft een beperkt beeld van de aanwezige stoffen in water. Vloeistofchromatografie met hoge resolutie massaspectrometrie (LC-ESI-HR-MS) combineert doelstof (target) en non-target analyse. Dit geeft een breed overzicht van de aanwezigheid stoffen in water [1, 2].

Bij non-target chemische screening van kwetsbare grondwaterwinningen van Vitens en het daaruit geproduceerd rein water is een scala aan onbekende stoffen aangetroffen. Het identificeren van een groot aantal onbekende stoffen is echter tijdrovend. Met behulp van datamining technieken trachten we meer duiding te geven aan deze onbekende stoffen en te prioriteren welke stoffen relevant zijn voor nader onderzoek.

Transformatieproducten vormen een wezenlijk deel van organische microverontreinigingen in de waterketen. Een literatuurstudie beargumenteerde dat transformatieproducten meer aandacht verdienen in de monitoring, evaluatie van zuiveringsprocessen, risicoanalyse en beleid [3, 4]. Transformatieprocessen (zoals oxidatie, hydrolyse en de-alkylering) leiden meestal tot meer polaire producten dan de moederstof. Bovendien ontstaan soms zogenaemde “dead end” transformatieproducten die onder de condities in het watersysteem of waterzuivering slecht verder afbreken en daardoor in concentraties worden aangetroffen die vergelijkbaar zijn met de moederstof.

1.2 Doel en onderzoeksvragen

Meetgegevens van Vitens laten zien dat er in verschillende drinkwaterbronnen (ruw) en het gezuiverde (rein) water duizenden onbekende features worden waargenomen. Dit zijn pieken (ion-counts) in het massa spectrum met een specifieke massa, retentietijd en fragmentatie.

Sommige features (stoffen) blijken aanwezig te zijn in zowel ruw als rein water, sommige stoffen verdwijnen tijdens het zuiveringsproces en andere stoffen verschijnen juist tijdens het zuiveringsproces.

Om meer duiding te geven aan deze features staan de volgende vragen centraal:

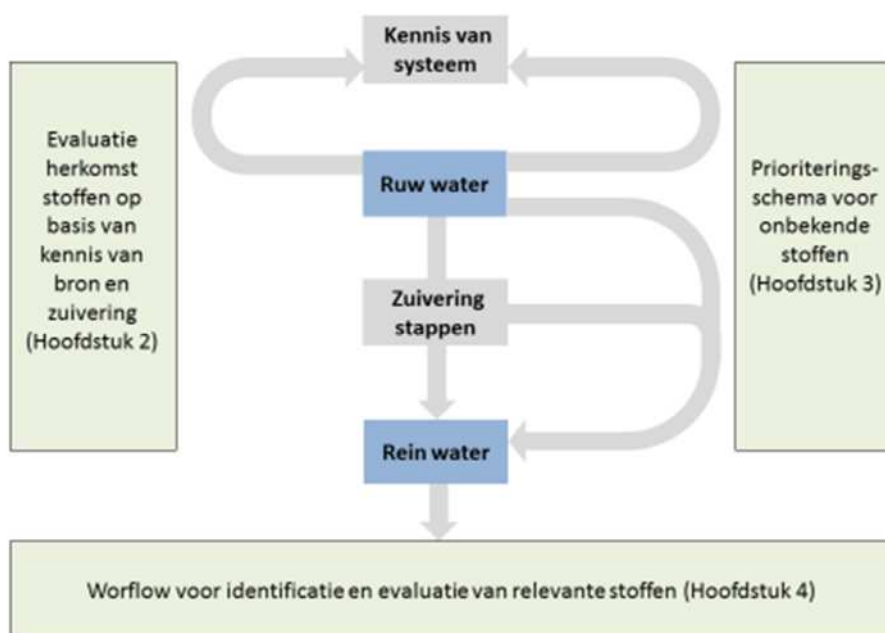
- Wat is de herkomst van de stoffen of stofgroepen: antropogeen of natuurlijk?
- Hebben de aangetroffen stoffen een relatie met het landgebruik en kan dat gebruikt worden voor hun prioritering?
- Nemen concentraties van stoffen toe of komen ze zelfs uitsluitend voor na de zuivering?
- Kunnen we transformatieproducten onderscheiden en aantonen?
- Hoe kunnen we onbekende stoffen (waaronder transformatieproducten) prioriteren voor identificatie en toxicologische evaluatie?
- Vormen de stoffen die in het reine water aanwezig zijn een risico voor de humane gezondheid?

1.3 Aanpak

De aanpak is onderverdeeld in drie activiteiten:

1. Evaluatie van de herkomst van features op basis van kennis van de bron correlaties met andere bekende antropogene stoffen en aanwezigheid van halogenen
2. Prioritering van features (respons, voorkomen, mogelijke herkomst, invloed zuivering, transformatieproduct)
3. Toepassing van de workflow voor identificatie en (eventuele) toxicologische duiding van geprioriteerde stoffen

De drie activiteiten zijn hieronder toegelicht.



Het doel van dit project is om een berekeningsmethode (tool) te formuleren waarmee in drinkwaterbronnen met non-target screening aangetroffen microverontreinigingen kunnen worden geprioriteerd en geïdentificeerd voor verdere risico evaluatie. Een groot aandeel van de aangetroffen stoffen betreft transformatieproducten, welke nog weinig in acht zijn genomen bij monitoring.

2 Duiding onbekende stoffen

2.1 Experimentele opzet en achtergrond

Als onderdeel van het SPO project 'Implementatie van LC-MS screeningstechnieken voor monitoring bedreigde winningen' zijn in 2014 29 winningen van Vitens in drie meetrondes gescreend met vloeistof chromatografie gekoppeld aan hoge resolutie massa spectrometrie (LC-HRMS) [5-7]. De selectie van 29 winningen is gebaseerd op een eerdere analyse van Vitens naar de hydrologische kwetsbaarheid (Tabel 2-1). De geselecteerde winningen zijn kwetsbare winningen, wat wil zeggen dat de kwaliteit van het water in de winningen wordt beïnvloed door bovengronds landgebruik of oppervlaktewater. Als gevolg van deze kwetsbaarheid zijn in veel van deze winningen regelmatig antropogene stoffen aangetroffen. In totaal zijn 131 individuele meetpunten in triplo bemonsterd en geanalyseerd. De meetpunten bestaan uit het reinwater, het gemengde ruwwater en een selectie van 2 of 3 pompputten per waterwinning. Van elke waterwinning is aan de hand van hydrologische modellen het landgebruik van het intrekgebied bepaald (Tabel 2-1).

De reversed phase (omgekeerde fase) chromatografie vindt plaats op een Waters, Atlantis-T3 kolom met een stationaire fase die zeer geschikt is voor polaire componenten. Het injectievolume is 1,0 ml water en de detectie geschied met een AB Sciex Q-TOF (API Triple TOF 5600+). De componenten worden in separate analysegangen in de positieve en negatieve ionisatie modus met accurate massa gemeten. Het geselecteerde massabereik van de detector (MS^1) is 100-1300 Da. Het massabereik van (MS^2) in de "high resolution" modus is 50-1300 Da. In de monsters worden duizenden onbekende features (stoffen) aangetroffen.

De gebruikte drempelwaarde van ion counts waarboven stoffen worden meegenomen is 500 voor de positieve en negatieve ionisatiemodus met een signaal ruis verhouding groter dan 5. Dit komt overeen met ongeveer 0,001 $\mu\text{g/L}$ Neburon equivalenten in zowel positieve als negatieve ionisatiemodus. Deze drempelwaarde is bewust vrij laag, om te voorkomen dat stoffen worden gemist. Ondanks goede voorzorgsmaatregelen kan in het traject van monsterneming tot en met analyse contaminatie optreden (blanco probleem). Dit kan leiden tot vals positieve resultaten. Alleen de features die een intensiteit hebben 5x hoger dan de procedure blanco zijn mee genomen in het onderzoek. Voor een goede analyse is enige chromatografische scheiding nodig. De injectiepiek ligt op 2,5 minuten. Het effluent tot 5 minuten wordt niet op de kolom gebracht. Zeer polaire stoffen met een retentie tijd tot 5 minuten, ongeveer 2,5 minuten na de injectiepiek passeren de kolom niet en worden derhalve niet geanalyseerd.

Deze features die voldoen aan de bovenstaande criteria kunnen individuele stoffen zijn maar in sommige gevallen kan één stof kan leiden tot meerdere features in het chromatogram. Verschillende features kunnen isotopen van elkaar zijn: hetzelfde molecuul met een verschillend aantal neutronen en elektronen in de kern van een of enkele atomen waardoor de totale massa van het molecuul versies met (net iets) verschillende massa's heeft. Daarnaast kunnen tijdens de analyse, en specifiek de ionisatie adducten worden gevormd, daarbij vormen stoffen en aanwezige ionen (veelal elementen zoals bijvoorbeeld natrium) samen een ion dat vervolgens wordt gedetecteerd met de massa van het molecuul en toegevoegde element. De features worden voor verder evaluatie opgeschoond van isotopen zodat de individuele stoffen zoveel mogelijk overblijven. Het verwijderen van adducten is echter niet altijd mogelijk omdat soms alleen het adduct wordt gemeten en niet de massa

van de moederstof zelf, als deze adducten worden verwijderd kan het betreffende molecuul niet of veel minder gevoelig gemeten worden. Om dataverlies te voorkomen is daarom gekozen adducten niet te verwijderen, maar bij latere interpretatie van resultaten na prioritering wel te onderzoeken of het een adduct betreft.

In totaal zijn 394 monsters zijn geanalyseerd. Deze monsters zijn afkomstig uit 29 winningen waarvan in de periode van een jaar 3 maal in ieder geval één ruwwatermonster en één reinwater monster is genomen.

TABEL 2-1 GESELECTEERDE WINNINGEN EN HET AANDEEL LANDGEBRUIK IN HET INTREKGEBIED (DE 25 JAARZONE). BRON: VITENS.

Winning	Aandeel bebouwd	Aandeel landbouw	Aandeel natuur
Amersfoort_Berg	62%	3%	35%
Beerschoten	16%	11%	73%
Bilthoven	59%	6%	35%
Corle	9%	74%	16%
De_Pol	17%	59%	22%
Druten	34%	57%	7%
Edese_Bos	6%	1%	93%
Ellecom	11%	13%	76%
Engelsewerk	38%	32%	24%
Epe	9%	19%	72%
Espelose_Broek	5%	80%	14%
Goor	34%	51%	13%
Groenekan	19%	66%	13%
Harderwijk	14%	1%	84%
Hasselo	62%	27%	7%
Havelterberg	7%	38%	55%
Heumensoord_I	57%	12%	31%
Heumensoord_II	21%	3%	76%
Holten	23%	36%	41%
Laren	29%	1%	69%
Losser	8%	60%	31%
Manderveen_NL	3%	60%	37%
Oldeholtgade	8%	70%	21%
Olden_Eibergen	9%	74%	15%
Schalteberg	10%	1%	89%
Soestduinen	9%	4%	87%
t'Klooster	5%	53%	42%
Wierden (2012)	30%	63%	7%
Zeist	70%	1%	29%
Zutphen	79%	10%	2%

Deze dataset is gebruikt om een prioriteringssysteem te ontwikkelen en te testen voor onbekende stoffen die worden waargenomen met behulp van non-target methoden. Het doel van de prioriteringsberekening (tool) is het omzetten van de meetgegevens van suspect stoffen in een score. Op basis van de score kunnen stoffen worden geselecteerd welke met de meeste urgentie in aanmerking komen voor verdere analyses.

2.2 Evaluatie herkomst stoffen ruwwater

Stoffen in het ruw- en drinkwater kunnen een diverse oorsprong hebben. Om de aangetroffen features beter te kunnen duiden, en uiteindelijk ook te prioriteren zijn diverse data-analyses uitgevoerd. Gebruikte basisgegevens zijn 11699 features gedetecteerd met LC-HRMS screening in positieve en negatieve ionisatie modus. Deze data bevat zowel 612 bekende (geïdentificeerde) stoffen als 11087 ongeïdentificeerde features waarvan enkel een respons (ion counts), een accurate massa en een retentietijd beschikbaar is. Om de herkomst van deze onbekende features te kunnen evalueren zijn de volgende (statistische) analyses uitgevoerd:

- Correlatie met kenmerken van de winningen, zoals het percentage jong water, het percentage oppervlaktewaterinvloed, en het landgebruik van het intrekgebied (Tabel 2-2).
- Correlatie met voorkomen macroparameters, te weten EGV, TOC, DOC, Cl, NO₃, SO₄, OXV, NH₄, redox, pH en B (Tabel 2-2)
- Correlatie tussen bekende en onbekende stoffen
- Het verschil in voorkomen tussen kwetsbare en niet-kwetsbare winningen

Een correlatie tussen de individuele stoffen en de verschillende kenmerken is significant wanneer de p-waarde van de steekproefuitkomst kleiner is dan 0,05 ($p < 0.05$). Op basis van de gegevens is gekozen om stoffen met correlatiecoëfficiënt groter dan 0,25 ($R^2 > 0,25$) te selecteren. Tabel 2-2 geeft weer hoeveel stoffen significant correleren met de genoemde kenmerken met een correlatiecoëfficiënt hoger dan 0,25.

TABEL 2-2 RESULTATEN CORRELATIE TESTEN.

Kenmerken en landgebruik van winningen	Aantal features met een correlatie met de volgende parameters	
	pos	neg
Percentage water <25jr	0	0
Percentage water<100jr	1	0
Percentage oppervlaktewater	19	0
Percentage bebouwd	0	0
Percentage landbouw	0	0
Percentage natuur	0	0
EGV	2	0
TOC	1	1
DOC	2	0
Cl	0	0
NO ₃	1	3
SO ₄	2	0
OXV	2	0
B	0	0
NH ₄	9	8
Redox	0	0
pH	0	0

In de volgende paragrafen zijn de verschillen toetsen per onderdeel beschreven en resultaten geëvalueerd.

2.3 Correlatie met hydrologische kenmerken en landgebruik van de winningen

De waterwinningen zijn door Vitens gekarakteriseerd met behulp van hydrologische modellen en systeemanalyse. De volgende hydrologische en kenmerken en typen landgebruik worden onderscheiden:

- Percentage onttrekkingswater dat jonger is dan 25 jaar
- Percentage onttrekkingswater dat jonger is dan 100 jaar
- Percentage onttrekkingswater afkomstig van oppervlaktewater
- Percentage van het intrekgebied (25-jaarszone) gekarakteriseerd als stedelijk gebied
- Percentage van het intrekgebied (25-jaarszone) gekarakteriseerd als agrarisch gebied
- Percentage van het intrekgebied (25-jaarszone) gekarakteriseerd als natuurgebied

De relatief lage correlatiecoëfficiënt van 0,25 is gekozen omdat er relatief beperkte correlatie is gevonden tussen landgebruik en hydrologische kenmerken. Uit de gekozen significantiecriteria volgt dat één stof correleert met het percentage water jonger dan 100 jaar (Tabel 2-2) en 19 stoffen, allen gedetecteerd in de positieve ionisatie modus, correleren met het percentage water afkomstig van oppervlaktewater. De 19 stoffen uit Tabel 2-4 zijn waarschijnlijk afkomstig van infiltrerend oppervlaktewater. Geen van de stoffen correleren met het percentage water jonger dan 25 jaar of het landgebruik van het intrekgebied.

TABEL 2-3 STOF DIE CORRELEERT MET HET PERCENTAGE ONTTREKKINGSWATER JONGER DAN 100 JAAR.

pieknr	Stof	R ²	p-waarde
403 (pos)	185,1016 / 11,61	0,26	<0,01

TABEL 2-4 STOFFEN DIE CORRELEREN MET HET PERCENTAGE OPPERVLAKTEWATER VAN HET ONTTREKKINGSWATER.

pieknr	Stof	R ²	p-waarde
81 (pos)	Chlortoluron	0,25	<0,01
161 (pos)	Isoproturon	0,25	<0,01
240 (pos)	10,11-dihydroxycarbamazepine	0,27	<0,01
286 (pos)	Tri-glyme	0,26	<0,01
481 (pos)	154,1217 / 18,91	0,27	<0,01
500 (pos)	171,1115 / 11,64	0,25	<0,01
525 (pos)	186,1477 / 12,30	0,28	<0,01
546 (pos)	197,1135 / 10,13	0,26	<0,01
548 (pos)	198,1340 / 12,97	0,27	<0,01
583 (pos)	211,1424 / 12,30	0,26	<0,01
607 (pos)	218,1380 / 10,53	0,25	<0,01
613 (pos)	220,1322 / 13,90	0,25	<0,01
670 (pos)	243,1331 / 15,86	0,27	<0,01
710 (pos)	264,1797 / 12,68	0,25	<0,01
715 (pos)	266,1385 / 20,84	0,26	<0,01
749 (pos)	284,2062 / 15,72	0,26	<0,01
757 (pos)	295,0871 / 22,10	0,26	<0,01
829 (pos)	340,1965 / 11,50	0,28	<0,01
844 (pos)	368,1554 / 10,32	0,26	<0,01

De resultaten laten zien dat er een beperkte correlatie is gevonden tussen individuele stoffen en de hydrologische- en landgebruikskenmerken van de winningen, en dat het aantal stoffen waarvoor zo'n lineaire correlatie wordt gevonden zeer beperkt is (<0,02% van de features). Echter minder dan de helft van de features is vaker dan 3 keer aangetroffen en een nog veel kleiner deel is in 3 of meer winningen aangetroffen. Pas als features in verschillende winningen worden aangetroffen is het mogelijk om een correlatie tussen eigenschappen van de winningen en features te bepalen. Daarnaast wordt de correlatie waarschijnlijk beïnvloed door specifieke verontreinigingen per locatie waardoor geen lineaire relatie met (bijvoorbeeld) de fractie oppervlaktewater aanwezig is. Deze verontreinigingen zijn niet specifiek kenmerkend voor winningen met gelijke kenmerken en nemen niet lineair toe met bepaalde kenmerken. Bovendien kan het landgebruik in het verleden zijn veranderd, kennen veel winputten gemengd landgebruik in hun intrekgebied en hebben sommige stoffen verschillende emissieroutes waardoor geen duidelijke relatie met één type landgebruik aanwezig is. Bijvoorbeeld: bestrijdingsmiddelen kunnen zowel uitspoelen na gebruik in de landbouw als via oppervlaktewater in de bodem terecht komen. Afvalwater en grondwater onder stedelijk gebied kunnen ook veel verschillende bestrijdingsmiddelen te bevatten omdat deze middelen worden/werden toegepast in tuinen en parken. Al met al kennen deze stoffen dus verschillende emissieroutes die de relatie met het bijvoorbeeld landgebruik

“landbouw” vertroebelen. De resultaten van deze exercitie zijn niet verder gebruikt bij het prioriteren van onbekende stoffen voor identificatie (Hoofdstuk 3) vanwege de beperkte (cor)relaties tussen landgebruik en voorkomen van stoffen.

2.4 Correlatie met macroparameters van het gewonnen water

Naast de verbindingen gedetecteerd met de LC-MS screening worden diverse anorganische en organische macroparameters gemeten. We toetsen of de onbekende organische stoffen samen voorkomen met deze macroparameters. De toets is uitgevoerd de volgende parameters:

- Elektrisch geleidingsvermogen (EGV)
- Totaal organisch koolstof (TOC)
- Opgelost organisch koolstof (DOC)
- Chloride (Cl)
- Nitraat (NO₃)
- Sulfaat (SO₄)
- Oxidatievermogen (OXV)
- Ammonium (NH₄)
- Redox milieu
- Zuurgraad (pH)
- Boor (B)

Uit deze gekozen significantie criteria volgt dat twee stoffen correleren met EGV, twee stoffen met TOC en twee stoffen met DOC (Tabel 2-4). TOC en DOC zijn algemene parameters voor het totaal (opgelost) organisch koolstof. Dit kan zowel van natuurlijke als antropogene oorsprong zijn. Een correlatie van een verbinding met DOC of TOC in water rijk aan natuurlijke humusverbindingen (bijvoorbeeld in veengebieden) zou kunnen wijzen op een van nature voorkomende stof. Terwijl correlatie met organisch materiaal afkomstig van vuilstorten of infiltrerend afvalwater, juist duidt op een stof van antropogene oorsprong.

Vier stoffen correleren met nitraat en twee met sulfaat (Tabel 2-5); al deze stoffen correleren tevens met het oxidatievermogen (NO₃ + SO₄). Daarnaast correleren 17 stoffen met ammonium. Mogelijk zijn deze stoffen gerelateerd aan agrarisch landgebruik, omdat nitraat, ammonium en sulfaat bestanddelen zijn van kunstmest. Geen enkele stof correleert met chloride of boor concentraties, het redox milieu of de zuurgraad.

TABEL 2-5 STOFFEN DIE CORRELEREN MET HET VOORKOMEN VAN ANORGANISCHE PARAMETERS.

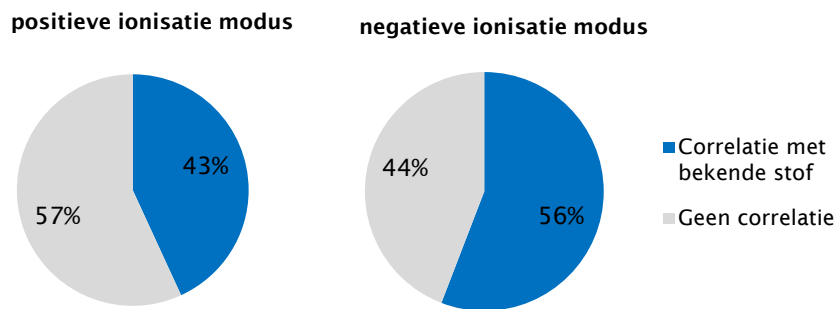
pieknr	stof	Parameter	R ²	p-waarde
665 (pos)	242,1244 / 11,49	EGV	0,26	P<0,05
319 (pos)	289,1178 / 10,43	EGV / DOC	0,27 / 0,29	P<0,05
1439 (pos)	286,0715 / 17,09	DOC	0,26	P<0,05
357 (pos)	473,1903 / 14,70	TOC / NH ₄	0,26 / 0,28	P<0,05
385 (neg)	305,0662 / 22,98	TOC / NH ₄	0,25 / 0,29	P<0,05
495 (neg)	257,1759 / 18,94	NO ₃ / OXV	0,25 / 0,29	P<0,05
533 (neg)	586,1685 / 14,71	NO ₃ / OXV	0,25 / 0,29	P<0,05
528 (neg)	378,9430 / 13,42	NO ₃ / OXV	0,26 / 0,3	P<0,05
428 (pos)	313,1092 / 13,85	NO ₃ / OXV	0,26 / 0,3	P<0,05
377 (pos)	185,1029 / 12,24	SO ₄ / OXV	0,39 / 0,35	P<0,05
403 (pos)	185,1016 / 11,61	SO ₄ / OXV	0,37 / 0,35	P<0,05
373 (neg)	267,1601 / 23,26	NH ₄	0,25	P<0,05
4920 (neg)	352,8520 / 23,88	NH ₄	0,26	P<0,05
4921 (neg)	352,8531 / 23,25	NH ₄	0,27	P<0,05
4922 (neg)	386,8127 / 26,05	NH ₄	0,27	P<0,05
4923 (neg)	388,8115 / 26,00	NH ₄	0,27	P<0,05
886 (neg)	443,2771 / 30,18	NH ₄	0,3	P<0,05
424 (neg)	251,0561 / 17,03	NH ₄	0,36	P<0,05
308 (pos)	228,1012 / 17,28	NH ₄	0,3	P<0,05
330 (pos)	328,0795 / 10,67	NH ₄	0,27	P<0,05
370 (pos)	485,1915 / 14,70	NH ₄	0,26	P<0,05
1252 (pos)	522,3063 / 23,00	NH ₄	0,27	P<0,05
1688 (pos)	261,1847 / 30,45	NH ₄	0,3	P<0,05
2074 (pos)	256,1184 / 13,83	NH ₄	0,27	P<0,05
4508 (pos)	399,2904 / 30,05	NH ₄	0,29	P<0,05
4515 (pos)	256,1188 / 14,48	NH ₄	0,26	P<0,05

De resultaten van deze exercitie werden vanwege de lage informatiedichtheid en het feit dat een gevonden correlatie vaak niet te vertalen is in een waarschijnlijkheid van antropogene of juist natuurlijke oorsprong niet gebruikt bij het prioriteren van onbekende stoffen voor identificatie (Hoofdstuk 3).

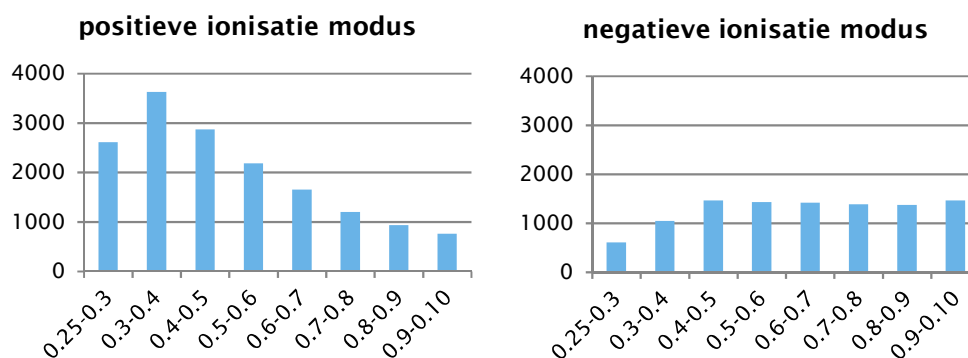
2.5 Correlatie bekende en onbekende stoffen

Het voorkomen van stoffen in combinatie met andere stoffen kan een aanwijzing zijn van de herkomst. Mogelijk hebben reeds geïdentificeerde stoffen en onbekende stoffen die sterk correleren een zelfde emissiebron en vergelijkbaar gedrag in de bodem.

De individuele stoffen en de specifieke kenmerken van de winningen vertoonden een correlatie wanneer er een significante correlatie is ($p < 0,05$) en de correlatiecoëfficiënt groter is dan 0,25. Uit deze criteria volgt dat de helft van de 11.087 features correleren met één of meerdere bekende antropogene stoffen. Stoffen gemeten in de negatieve ionisatie modus correleerden vaker dan stoffen gedetecteerd in de positieve ionisatie modus (Figuur 2-1). De verdeling van de bijbehorende correlatie coëfficiënten zijn weergegeven in Figuur 2-2.

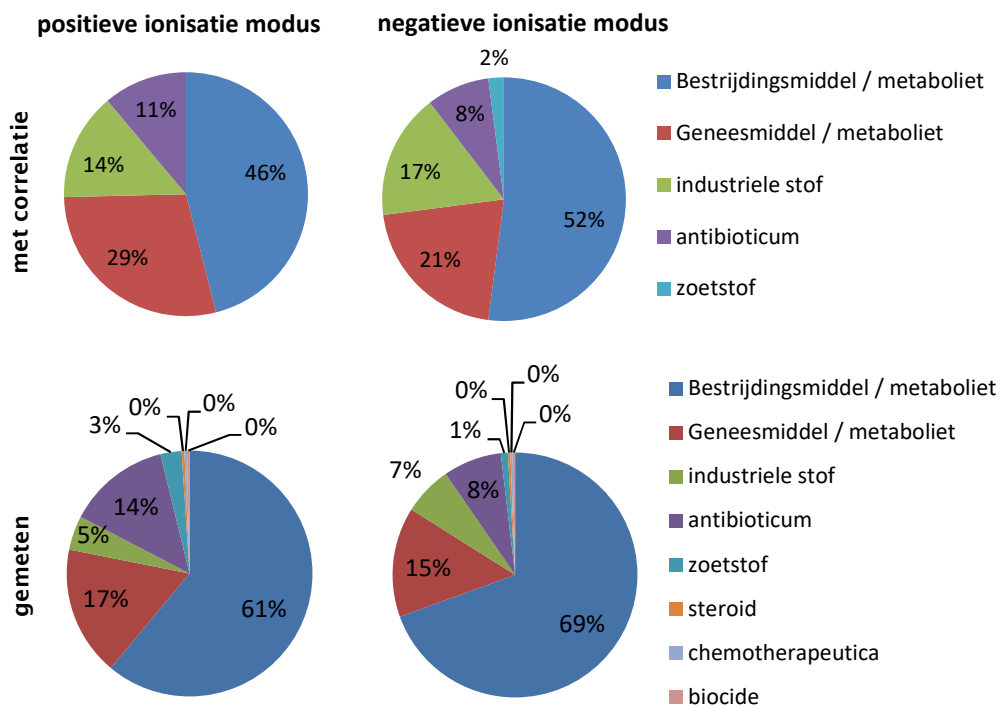


FIGUUR 2-1 VERDELING ONBEKENDE STOFFEN MET WEL/GEEN CORRELATIE MET ÉÉN OF MEER BEKENDE ANTROPOGENE STOFFEN.



FIGUUR 2-2 VERDELING VAN HET AANTAL SIGNIFICANTE ($P < 0,05$) CORRELATIES TUSSEN ONBEKENDE STOFFEN EN BEKENDE ANTROPOGENE STOFFEN GESORTEERD OP MATE VAN CORRELATIE (MET EEN CORRELATIE COEFFICIENT GROTER DAN 0,25).

Van de 612 bekende stoffen correleerden 99 stoffen met één of meer van de aangetroffen onbekende stoffen. Deze bekende 99 stoffen zijn onder te verdelen in bestrijdingsmiddelen, geneesmiddelen, waaronder antibiotica en industriële stoffen (Figuur 2-3). De meeste onbekende stoffen correleren met een bestrijdingsmiddel, gevolgd door een geneesmiddel, een antibioticum en een industriële stof (Tabel 2-6). Dat is mogelijk het gevolg van meer meetgegevens voor bestrijdingsmiddelen en meer bestrijdingsmiddelen in het meetpakket. Echter, de verdeling van de correlaties van onbekende stoffen over de vier stofgroepen is niet evenredig ($\chi^2 = < 0,01$). Figuur 2-3 laat zien van alle correlerende stoffen een kleiner aandeel wordt ingedeeld in de groep bestrijdingsmiddelen dan het aandeel bestrijdingsmiddelen van alle stoffen dat werd gemeten. Daarentegen is het aandeel van de geneesmiddelen en industriële stoffen van alle correlerende stoffen groter dan het aandeel geneesmiddelen en industriële stoffen van alle gemeten stoffen. Voor de antibiotica ligt het aandeel aangetroffen en gemeten van alle stoffen ongeveer gelijk. De conclusie is dat er meer correlaties zijn tussen features en geneesmiddelen dan je zou verwachten op basis van het meetpakket van target stoffen; voor bestrijdingsmiddelen zijn er minder correlaties.



FIGUUR 2-3 VERDELING BEKENE STOFFEN IN STOFGEROEPEN DIE CORRELEREN MET ÉÉN OF MEER ONBEKENE STOFFEN (BOVEN) OF ZIJN GEMETEN (ONDER).

TABEL 2-6 AANTAL CORRELATIES MET BEKENE EN ONBEKENE STOFFEN, PER STOFGROEP.

Stofgroep	Aantal correlaties met onbekende features
bestrijdingsmiddelen / metabolieten	14.573
geneesmiddelen / metabolieten	9.789
antibiotica	3.321
industriële stoffen	3.071
zoetstof	41
Totaal	30.795

De resultaten van deze exercitie werden gebruikt bij de prioritering van onbekende stoffen voor identificatie (Hoofdstuk 3). De correlatie van een onbekende stof met een bekende (antropogene) stof duidt op een waarschijnlijke antropogene oorsprong. In paragraaf 3.3.4 wordt de scoreberekening weergegeven voor antropogene oorsprong als onderdeel van de prioritering.

2.6 Vergelijk voorkomen in kwetsbare en niet kwetsbare winningen

Om na te gaan of features (stoffen) van nature voorkomen in het ruwwater of het resultaat zijn van antropogene beïnvloeding, zijn alle winningen van Vitens ingedeeld naar kwetsbaarheid. Deze indeling is gebaseerd op de reistijdverdeling van het onttrokken water en gegevens over het voorkomen van antropogene verontreinigingen [8]. Winningen met een relatief groot aandeel van water met korte reistijden en/of aanwezigheid van antropogene stoffen is als kwetsbaar aangemerkt. De locaties waar het water op basis van modellen duidelijk al erg lang (minimaal meerdere honderden jaren) in grond zit voor het oppervlaktewater of ondiep grondwater in de put komt, worden aangemerkt als niet kwetsbaar. De categorie ertussen wordt aangemerkt als matig kwetsbaar. De 29 geselecteerde winningen zijn allen aangemerkt als kwetsbaar. In deze analyse zijn echter ook gegevens van 2016 andere winningen uit een meetserie van 2016 meegenomen. Van deze winningen zijn 75 als niet kwetsbaar, 14 als matig kwetsbaar en 22 als kwetsbaar aangemerkt. Op deze manier kon worden nagegaan of er in de analyses toch bekende antropogene stoffen in de niet kwetsbare locaties worden aangetroffen. Dit zou kunnen wijzen op kortsluitstromen in de ondergrond waardoor toch recent antropogeen beïnvloed water in de put kan komen waarvan wordt verondersteld wordt dat het water veel ouder is. Deze locaties zijn vervolgens aangemerkt als matig kwetsbare locaties. Stoffen die alleen voorkomen in kwetsbare winningen (waarschijnlijk antropogeen beïnvloed) zijn mogelijk van antropogene oorsprong. Tevens konden op deze manier features die voornamelijk voorkomen in niet kwetsbare winningen aangemerkt worden als mogelijk van natuurlijke oorsprong. Deze zijn door natuurlijke processen in de ondergrond zijn gevormd. Features die uitsluitend voorkomen in niet kwetsbare winningen vormen een bijzondere groep: ze zijn mogelijk kenmerkend zijn voor dat type winningen. Ze worden uiteraard ingedeeld als een waarschijnlijk natuurlijke stof. Van veel voorkomende features kan worden verwacht dat ze in alle winningen voorkomen, zowel kwetsbaar als niet kwetsbaar. De features die alleen voorkomen in kwetsbare winningen, dus winningen die waarschijnlijk antropogeen beïnvloed zijn, worden ingedeeld in de groep mogelijk antropogeen.

Op basis van de feature lijst van meetgegevens van ruwwater analyses uit 2016 en het percentage voorkomen in bedreigde en onbedreigde winningen is een indeling gemaakt van natuurlijke en antropogene stoffen. In de analyse zijn de kwetsbare en matig kwetsbare winputten samengevoegd tot bedreigde winputten. De criteria van de stofindeling zijn weergegeven in Tabel 2-7.

TABEL 2-7 CRITERIA INDELING STOFFEN

Stofgroep	Criterium stofgroep
Waarschijnlijk natuurlijke stof	Stof komt >30% voor in onbedreigd en in bedreigd of Stof komt alleen voor in onbedreigd
Mogelijk natuurlijke stof	Stof komt >10% voor in onbedreigd en in bedreigd
Misschien natuurlijke stof	Stof komt <10% voor in onbedreigd en in bedreigd
Mogelijk antropogeen	Stof komt alleen voor in bedreigde winningen
Komt weinig voor	Wordt minder dan 3 maal aangetroffen

Uit de dataset van 2016 zijn van de features die 3 keer of vaker zijn aangetroffen, in totaal 39 features (positief en negatief) aangemerkt als 'waarschijnlijk natuurlijk'. 324 features zijn mogelijk natuurlijk en 392 misschien natuurlijk. Veruit het grootste deel van de features die meer dan 3 keer zijn aangetroffen vallen in de categorie 'mogelijk antropogeen' (3307 features=81%). De resultaten van deze exercitie werden gebruikt bij de prioritering van onbekende stoffen voor identificatie (Hoofdstuk 3).

2.7 Chemische samenstelling van onbekende stoffen

Het voorkomen van Cl en Br in stoffen kan een aanwijzing zijn van antropogene herkomst [9]. Hiervoor wordt de brutoformule gebruikt als deze bekend is, of het isotopenpatroon van een feature. Cl en Br hebben een karakteristiek isotopenpatroon, en zijn dus eenvoudig herkenbaar (Tabel 2-8) in de features. Vitens heeft gezocht op isotopen aan de hand van de twee Da massa verschil waarbij deze in hetzelfde monster aanwezig moeten zijn. De resultaten staan in de werkbladen S_IsoAdd en Iso_AddNaamOud. Gebaseerd daarop kunnen Cl en Br houdende features in de prioritering opgenomen worden, ook als een brutoformule ontbreekt.

TABEL 2-8 ELEMENTEN EN HUN ISOTOPENPATRONEN.

Element	Massa (Da)	+1 isotoop (Da)	%	Vershil	+2 isotoop (Da)	%	Vershil
H	1,00783	2	0	1			
C	12	13	1,1	1			
N	14,0031	15	0,4	1			
O	15,9949	17	0	1	18	0,2	2
S	31,9721	33	0,8	1	34	4,5	2
Cl	34,9689				37	32	2
Br	78,9183				81	97	2

De resultaten van deze exercitie werden gebruikt bij de prioritering van onbekende stoffen voor identificatie (Hoofdstuk 3). Stoffen met één of meerdere halogeen verbindingen zijn waarschijnlijk van antropogene oorsprong omdat gehalogeneerde organische stoffen in de natuur maar beperkt voorkomen. De resultaten vormen deel van de scoreberekening voor antropogene oorsprong in paragraaf 3.3.4.

2.8 Verwijderingsefficiëntie

Tijdens de zuivering kunnen stoffen worden verwijderd of gevormd. Hiervoor berekenen we voor elke onbekende stof de ratio tussen de concentratie in het ruwe water en het reine water na zuivering op de drie tijdstippen van monsternamen. Stoffen zijn niet altijd op alle drie de meetmomenten waargenomen. Pas wanneer de ratio kleiner is dan 0,75 of de ratio groter is dan 1,25 op minstens twee van de drie tijdstippen per locatie gaan we respectievelijk uit van af- of toename. Als de stof alleen aanwezig is in ruwe water of reine water wordt de stof respectievelijk verwijderd of gevormd. Er is bewust gekozen voor deze relatief arbitraire grens omdat statistische analyse van af of toename door analyse van verhoudingen tussen ruw en rein water niet altijd goed mogelijk was door beperkte hoeveelheid gegevens per feature. Wanneer er genoeg data beschikbaar is maar niet wordt

voldaan aan bovenstaande criteria veronderstellen we geen vorming of verwijdering. De feature 9stof0 is dan stabiel tijdens behandeling. De resultaten van deze exercitie werden gebruikt bij de prioritering van onbekende stoffen voor identificatie (Hoofdstuk 3). In paragraaf 3.3.3 wordt de scoreberekening weergegeven voor de verwijdering of vorming van stoffen als onderdeel van de prioritering. Indien de concentratie tijdens de zuivering toeneemt of wordt gevormd draagt dit bij aan een hogere prioritering van de onbekende stof, wanneer deze afneemt of verdwijnt, draagt dit bij aan een lagere prioritering.

2.9 Transformatieproducten

Transformatieproducten (TP) ontstaan in het milieu en tijdens de waterbehandeling en kunnen een relevante fractie van de organische microverontreinigingen in de waterketen vormen. Omdat maar een deel van de transformatieprocessen bekend is, en de TP die daaruit voortkomen veelal onbekend zijn, zijn ze veelal niet routinematig opgenomen in de monitoring [10]. Om het voorkomen van TP in kaart te brengen heeft Vitens in de features van de LC-HRMS non-target screening data gezocht op paren met bepaalde massaverschillen, die mogelijk een transformatieproces aantonen. Aan hand van de lijst met transformatieprocessen en de daardoor ontstane massaverschillen tussen moederstoffen en TP (

Tabel 2-9, gebaseerd op [11]) worden features gelabeld als potentiële moederstoffen of TP als deze in 'ruw' en 'rein' waterkoppels aanwezig zijn (zie bijgeleverde Excel) dit betreffende onderdeel is beschreven in werkblad S_TP). TP worden per definitie gevormd uit moederstoffen, wat inhoudt dat hun signaal in LC-HRMS-experimenten met de tijd toeneemt, d.w.z. tussen 'ruw' en 'rein' monsters, terwijl dat van hun moederstoffen afneemt. Deze signaalintensiteitspatronen worden gebruikt bij trendfiltering. De hier gebruikte criteria waren ratio moederstof (rein/ruw) $< 1,1$ (afnemen), ratio TP (rein/ruw) $> 0,9$ (toenemen). Verder worden criteria aangaande massa-afwijking van het massaverschil ($< 0,0025$ Da, ie ~ 5 ppm bij 500Da), en de retentietijd van de TP - afhankelijk van het transformatieproces hoger of lager dan de RT van de moederstof - toegepast (zie Tabel 2-10). Overigens kan een feature een TP zijn van meerdere moeders, maar een ander soort TP.

TABEL 2-9. TRANSFORMATIEPROCESSEN EN DAARDOOR ONTSTAAN MASSA VERSCHILLEN TUSSEN MOEDERSTOFFEN EN TP.

Transformatie	Atoom- verlies / - winst	Formule- verschil	Massaverschil	RT van TP
Oxidatie	+	O	15,99490	lager
Oxidatie	+	O2	31,98980	lager
Oxidatie	+	O3	47,98470	lager
Oxidatie	+	O4	63,97960	lager
Demethylering	-	CH2	-14,0157	lager
Deethylering	-	C2H4	-28,03130	lager
Dehydratie	-	H2O / Lactam	-18,01060	lager
Chloor reductie	-	Cl/+H	-33,96110	lager
Acetylering	+	C2H2O	42,01060	lager
Deacetylering	-	C2H2O	-42,01060	hoger
Glucuronidatie	+	C6H8O6	176,03200	lager
Deglucuronidatie	-	C8H8O6	-176,03200	hoger
Sulfonatie	+	SO3	79,95680	lager
Desulfonatie	-	SO3	-79,95680	hoger
Anthranilacid isopropylamide	-	+2H-SO2	-61,94630	hoger
DesIsoPropyl	-	-C3H3	-42,04700	lager
Hydroxylering	-	(-Cl, +OH)	-17,96614	lager
Desamino (-HN)	-	(-HN)	-15,01090	lager
Dimethenamid (OA)	-	OA	-3,98687	lager
Dimethenamid (ESA)	-	ESA	45,99579	lager

Deze aanpak gebaseerd op de zogenoemde “metabolic logic” leverde veel hits op [11]. Omdat er meer dan 11.000 features zijn bestudeerd bestaat de kans dat toevallig features worden gevonden met bovenstaande massaverschillen. In de toekomst kunnen extra zoekregels nodig zijn om transformatieproducten van toevallige hits te scheiden. Er kan bijvoorbeeld worden overwogen welke transformatieprocessen bij welke zuiveringstechnieken optreden. Vervolgens kan voor zandfiltratie, beluchting, pH aanpassing, ontharding en actief kool met geselecteerde transformatieprocessen naar TP gezocht worden. Ook kunnen de ratio-criteria strenger worden gemaakt, en statistische analyses toegepast worden om vals positieven te reduceren. Voornamelijk significantietests gebaseerd op Student's t-test en fold changes kunnen features aantonen die significant verschillen tussen ‘ruw’ en ‘rein’ koppels. Aan hand van vulcano plots kunnen die resultaten effectief gevisualiseerd worden [12]. Bovendien kunnen multivariate analyse (MVA) methoden, zoals principal component analyse (PCA), hiërarchische clustering en partial least squares analyse (PLS), worden gebruikt voor het karakteriseren en groeperen van co-variërende features. Zo is MVA met succes toegepast om non-target features te classificeren in moederstoffen en TP [11]. Deze methoden vielen echter buiten de scope van dit SPO onderzoek.

Een andere reeks strategieën voor het identificeren van TP maakt gebruik van de informatie die inherent is aan een verbinding, dat wil zeggen de massadefecten, isotooppatronen en structurele samenstelling. Deze veronderstelt dat de moederstoffen en TP-informatie in zekere mate gelijk zijn, of vergelijkbaar. In massadefectfiltering, wordt het verschil tussen de nauwkeurige en de nominale massa van een ion, het zogenaamde massadefect, gebruikt om

multi-isotoopelementen, in het bijzonder halogenen, te identificeren [13] Dit kan TP-identificatie vergemakkelijken door kenmerken met specifieke functionele groepen of elementen bloot te leggen. Evenzo kunnen isotopenpatroonfilters worden gebruikt om TP te identificeren op basis van de isotoopatronen van hun moederstof [14]. Ten slotte kan de structurele gelijkenis van moederstoffen en hun TP onthullen; omdat de structuren vergelijkbaar zijn, is ook de stabiliteit of reactiviteit van bepaalde delen van het molecuul vergelijkbaar en vertonen deze moleculen gemeenschappelijke karakteristieke fragmentionen. Een interessante aanpak voor toekomstige TP-bevestiging is dus een criterium gebaseerd op de mate van de gelijkenis tussen MS² fragmentatiespectra tussen moederstof en potentiële TP, zoals recent beschreven [15]. Dit werk toont aan dat de gelijkenis van een MS² fragmentatiespectrum op een structurele gelijkenis kan duiden, en dat dit kan worden gebruikt voor het screenen voor verwante verbindingen, zoals TP.

2.10 Van stoffeigenschap naar duiding

LC-MS analyses geven een breed beeld van de waterkwaliteit, maar het duiden en identificeren van stoffen is tijdrovend. Met behulp van een aantal datamining technieken is getracht meer duiding te geven aan gedetecteerde features. Landgebruik en hydrologische kenmerken van de waterwinning blijken een lastige voorspeller te zijn voor het voorkomen van bepaalde features in het ruwe water en zijn niet meegenomen in de uiteindelijke prioritering van deze features (Hoofdstuk 3). Echter, we hebben enkele relaties kunnen ontdekken tussen features en hydrologische- of landgebruikskennmerken. 19 features zijn waarschijnlijk afkomstig van infiltrerend oppervlaktewater, omdat ze een correlatie vertonen met het percentage oppervlaktewater in het onttrekkingswater. Daarnaast zijn 23 stoffen waarschijnlijk afkomstig zijn van agrarische toepassingen, omdat de stoffen samen voorkomen bestanddelen van kunstmest (NO₃, SO₄, NH₄). Vier stoffen daarentegen correleren met algemene parameters TOC en DOC, wat zou kunnen duiden op een bestanddeel van natuurlijk organisch materiaal.

De waarschijnlijkheid dat een feature van antropogene oorsprong is, is meegenomen in de prioritering, zie de scoreberekening voor de evaluatie van de antropogene oorsprong in paragraaf 3.3.4. De volgende drie methoden leveren waardevolle informatie over de waarschijnlijkheid van de antropogene oorsprong van de onbekende stoffen:

- *Onderlinge correlaties tussen bekende antropogene en onbekende stoffen.* Onbekende stoffen die vaak samen voorkomen met bekende antropogene stoffen zijn waarschijnlijk zelf ook antropogeen. Voor de helft van de features is een correlatie gevonden met een bekende stof; met 99 van de 612 doelstoffen.
- *Voorkomen van stoffen in kwetsbare (beïnvloed door antropogene verontreinigingen/activiteiten) niet-kwetsbare (niet beïnvloed door antropogene verontreinigingen/activiteiten) winningen.* Het vaker voorkomen van bepaalde onbekende stoffen in onbedreigde niet-kwetsbare winningen kan duiden op natuurlijke oorsprong (bijvoorbeeld afkomstig van humus).
- *De aanwezigheid van halogenen in onbekende stoffen.* Door industrie geproduceerde stoffen bevatten vaak(er) halogenen dan natuurlijke stoffen. Industrie maakt organische verbindingen met halogenen om ze specifieke eigenschappen te geven zoals het bevorderen van hun stabiliteit in organismen of het milieu.

3 Prioriteren van onbekende stoffen

3.1 Aanpak

Het prioriteren van onbekende features in ruw en rein water is ingewikkeld omdat er beperkte informatie is van de structuur en eigenschappen van de stoffen. Doordat er zeer veel onbekende features aangetroffen zijn, is het onmogelijk om deze features allemaal in detail te onderzoeken en de stoffen die leiden tot de waarneming van deze features te identificeren. Het is noodzakelijk om de onbekende features te prioriteren en de features met de hoogste prioriteit eerst aan nader onderzoek te onderwerpen. Op basis van de prioritering worden features geselecteerd voor chemisch-analytische bevestiging. Dit proces wordt beschreven in Hoofdstuk 4. Vervolgens kan voor een deel van de vermoedelijke of geïdentificeerde stoffen een uitgebreide gezondheidskundige risico evaluatie worden uitgevoerd met behulp van toxicologische informatie en aangetroffen concentraties. Deze evaluatie is niet in het voorliggende rapport uitgevoerd.

Bij de prioritering worden de volgende aspecten meegenomen:

- 1) De relatieve intensiteit genormaliseerd naar een interne standaard van een feature met een unieke massa en retentietijd wordt gekwantificeerd in de met de subscore "TotalIntensity". Features die isotopen zijn van andere features worden (zoveel mogelijk) verwijderd door middel van het uitsluitcriterium "Nolsotope".
- 2) De frequentie van aantreffen in een serie van monsters wordt gekwantificeerd in de met de subscore "TotalFrequency". De frequentie van aantreffen zegt ook iets over de betrouwbaarheid van een waarneming. Het uitsluitingscriterium "TotalRepeat" geeft de absolute ondergrens van het aantal waarnemingen aan waaronder een stof niet meer relevant wordt geacht aan.
- 3) Zuiveringsprocessen kunnen verschillende effecten hebben op de aanwezigheid van stoffen. De hoeveelheid van een stof kan verminderen (verwijdering), niet veranderen, of stijgen (vorming van transformatieproducten¹ of worden toegevoegd door materialen gebruikt in de zuivering). De verandering van de intensiteit van een feature wordt gekwantificeerd met de subscore "TotalChange".
- 4) De mogelijke antropogene herkomst van een feature wordt gewogen met behulp van de subscore "TotalAntropogenic". Features met een (mogelijk) antropogene herkomst krijgen een hogere score dan features met een (waarschijnlijk) natuurlijke oorsprong omdat de aanwezigheid van antropogene features antropogene beïnvloeding van een bron indiceren. Bovendien bestaat voor stoffen van antropogene stoffen een beter handelingsperspectief (emissiereductie) dan voor natuurlijke stoffen.

In de paragrafen hieronder worden de verschillende componenten van de prioritering besproken en de prioritering toegelicht.

¹ Transformatieproducten worden gevormd in chemische reacties in zuiveringsprocessen. Deze stoffen zijn vaak meer polaire producten van de moederstof, en kunnen zowel meer of minder reactief en/of toxisch zijn. Er zijn ook transformatieproducten welke slecht afbreken en daardoor meer persistent zijn dan de moederstof.

3.2 Data voorbereiding

3.2.1 Normaliseren van data

Non target screening richt zich op het meten van een breed spectrum van organische verbindingen. Of deze verbindingen worden aangetroffen door de methode wordt bepaald door de concentratie van de stof maar ook door stofeigenschappen, omstandigheden en materialen en technieken toegepast van bemonstering, transport, opslag, scheiding, detectie en data analyse. Hieronder zijn deze factoren toegelicht.

- Het gedrag van de stof gedurende bemonstering, transport en opslag is afhankelijk van de vluchtigheid, stabiliteit van de stof onder de condities tijdens dit proces (bijv. schommelingen in Redox condities, temperatuur, pH, zoutgehalte, licht en samenstelling maar ook contact met gebruikte monstermaterialen)
- Het gedrag van de stof gedurende monstervoorbehandeling is afhankelijk van toegepaste voorbehandelingen. Vitens past directe injectie toe om invloed van voorbehandeling zo veel mogelijk te beperken.
- Het gedrag van de stof tijdens chromatografische scheiding. Reversed phase vloeistof chromatografie kan een breed scala aan stoffen scheiden. Zeer polaire stoffen worden met deze methode echter niet of beperkt tegengehouden waardoor ze met de injectiepiek elueren [16].
- Voor de detectie van de stoffen moeten de stoffen worden geïoniseerd. De combinatie van de ionisatietechniek, chemische structuur en samenstelling van het eluent bepaalt in hoeverre een stof kan worden geïoniseerd. (Gehalogeneerde) koolwaterstoffen worden niet of nauwelijks geïoniseerd terwijl stoffen met allerlei actieve groepen (bijv. met zuurstof, stikstof, zwavel) makkelijk(er) worden geïoniseerd. Het gevolg hiervan is dat sommige stoffen met zeer hoge concentraties niet te meten zijn terwijl andere stoffen met orde grootte lagere concentraties wel meetbaar zijn [17, 18].

Om de ion counts te normaliseren worden alle responsen van onbekende stoffen gecorrigeerd voor de interne standaard Neburon. Deze interne standaard wordt zowel in de positieve als negatieve ionisatiemodus toegepast. De respons van neburon is in de negatieve ionisatie modus lager ten opzichte van de gemiddelde totale respons van alle features dan in de positieve ionisatiemodus. De gemiddelde respons van de positieve monsters was 4,9 µg/L equivalenten van Neburon, voor de negatieve ionisatiemodus is dit 18.6 µg/L equivalenten van Neburon. Mogelijk is dit een gevolg van een groter aantal natuurlijke organische zuren (humus bestanddelen) die in de negatieve ionisatiemodus een grotere bijdrage leveren aan de respons. Om de positieve en negatieve ionisatie een gelijk gewicht te geven in verhouding met andere componenten van de prioritering zijn de responsen door de desbetreffende getallen gedeeld zodat de uiteindelijke score van de intensiteit in positieve en negatieve modes een (ongeveer) gelijke score krijgen. Op de prioritering van positief en negatief geïoniseerde features heeft dit overigens geen invloed omdat in beide modi een aparte prioriteringslijst is gemaakt.

3.2.2 Adducten en Isotopen

LC-HRMS non-target screening van watermonsters genereert duizenden features die een accurate massa en retentietijd (RT) vertegenwoordigen. De hoeveelheid van features voorkomt echter vaak de identificatie van verbindingen. Strategieën om het aantal features te reduceren worden dus toegepast. Deze kunnen door de vermindering van de gegevenscomplexiteit leiden tot hogere fractie identificaties (identification rates) met lagere vals positieven (false discovery rates). Een reductiestrategie is het groeperen van features die van dezelfde verbinding afstammen, d.w.z. monoisotopische piek plus isotopen, adducten - inclusief

complexe adducten tussen features en achtergrondpieken, en homodimeren, heterodimeren en hogere n-meer en / of in-source fragmenten, samenvattend ook genoemd “degenererde pieken” [19]. Onderzoek heeft aangetoond dat MS-analyse van een enkele stof kan leiden tot detectie van >100 degenererde pieken. Het is dus nodig deze pieken te annoteren en mogelijk vervolgens te verwijderen met behoud van die welke overeenkomen met unieke verbindingen. Daartoe heeft Vitens gezocht op isotopen aan de hand van de massa verschillen in Tabel 2-7 en Na, NH₄, K, CH₃OH, ACN, en isopropanol adducten, waarbij deze in hetzelfde monster aanwezig moeten zijn. De resultaten staan in de werkbladen S_IsoAdd. Gebaseerd daarop kunnen mono-isotopische pieken in de prioritering worden opgenomen, en de anderen isotopische pieken en adducten uit de feature lijst verwijderd.

3.3 Prioriteren van onbekende stoffen

De onbekende stoffen zijn geprioriteerd met de onderstaande criteria.

3.3.1 Subscore Total_Intensity

De subscore voor de intensiteit wordt bepaald door de logaritmisches getransformeerde waarde van interne standaard gecorrigeerde maximale responsen van een feature of in rein water. Deze waarden zijn logaritmisches getransformeerd omdat de verdeling van de absolute waarden geen normale verdeling liet zien ($P < 0.001$ voor de positieve en negatieve ionisatie (Kolmogorov-Smirnov test, Graphpad Prism 5.0)). Vervolgens zijn deze waarden genormaliseerd tot een positieve waarde in verband met verdere berekeningen door de feature met de laagste maximale respons op 0 te stellen en de rest navenant op te hogen. In deze studie is gekeken naar stoffen die (ook) in rein water zijn aangetroffen omdat dit van groter belang wordt geacht dan het aantreffen van stoffen in het ruwe water alleen. In principe kan dit worden aangepast afhankelijk van het doel van de prioritering (bronnenonderzoek, of waterkwaliteit).

$$TotalIntensity = \log(\max intensity\ rein) + normalisatiewaarde$$

$$normalisatiewaarde = -(\log(\max intensity\ rein\ of\ feature\ with\ lowest\ max\ response))$$

3.3.2 Subscore Total_Frequency

De subscore voor de frequentie wordt bepaald door de logaritmisches getransformeerde waarde van het aantal waarnemingen van een feature in rein water monsters. Deze waarden zijn logaritmisches getransformeerd omdat de verdeling van de absolute waarden geen normale verdeling liet zien ($P < 0.001$ voor de positieve en negatieve ionisatie (Kolmogorov-Smirnov test, Graphpad Prism 5.0)). Er wordt gekeken naar rein water omdat de aanwezigheid van stoffen in het reine water in deze studie van groter belang wordt geacht dan het aantreffen van stoffen in het ruwe water. In principe kan dit worden aangepast afhankelijk van het doel van de prioritering (bronnenonderzoek, of waterkwaliteit).

$$TotalFrequency = \log(n\ rein)$$

3.3.3 Subscore Total_Change

De subscore voor de verandering van de respons als gevolg van zuivering worden per winning gegevens vergeleken tussen monsters van het ruw en rein water. Voor elke winning waar de feature toeneemt (>25%) krijgt de stof een hogere score, en voor elke plek waar de stof afneemt (>25%) een lagere score. Voor alle winningen definiëren we of een onbekende stof (gedeeltelijk) wordt verwijderd of gevormd. Vervolgens leiden de waarden voor de individuele winningen samen tot één waarde voor de betreffende feature. Tabel 3-1 geeft een overzicht van de (gedeeltelijk) gevormde en verwijderde stoffen.

TABEL 3-1 VERWIJDERING OF VORMING VAN STOFFEN TIJDENS ZUIVERING PER WINNING

Definitie gedrag zuivering	Omschrijving	Positieve ionisatie modus	Negatieve ionisatie modus
Concentratie neemt af ...waarvan stof wordt geheel verwijderd*	Afname >25% op minstens twee tijdstippen	1320	1325
	Stof wordt verwijderd op minstens twee tijdstippen	1059	1211
Concentratie neemt toe ...waarvan stof ontstaat tijdens zuivering*	Toename >25% op minstens twee tijdstippen	317	160
	Stof wordt gevormd op minstens twee tijdstippen	216	94
Concentratie blijft gelijk	Afname of toename <25%	599	322

*Maakt deel uit van bovenstaande

Alle scores worden op de volgende wijze per feature opgeteld en gedeeld door het aantal zuiveringen waar de stof is aangetroffen.

$$TotalChange = \frac{0,5 * n_{afname} + 1,5 * n_{toename} + 1,0 * n_{ongewijzigd}}{aantal\ waarnemingen}$$

Dit levert een getal op tussen de 0,5 en 1,5. Er wordt in de score geen onderscheid gemaakt tussen gedeeltelijke verwijdering of vorming en volledige verwijdering of vorming omdat dit o.a. afhankelijk is van de intensiteit van de waarneming ten opzichte van de meetgevoeligheid. Features die zijn waargenomen met een lage intensiteit worden namelijk eerder aangemerkt als gevormd of verwijderd, terwijl features met hoge intensiteit voldoende resolutie hebben om af of toename te registreren. In de onderzochte winningen worden typen zuiveringen en productieprocessen toegepast, welke steeds een andere invloed hebben op de aanwezigheid van microverontreinigingen en transformatieproducten daarvan. Dit zijn:

- *Zandfiltratie.* In water aanwezige deeltjes worden afgevangen in een zandlaag, waarbij verdere zuivering plaats kan vinden in fysisch-chemische en biologische processen in de zandlaag.
- *Beluchting.* Beluchting heeft verschillende doelen. Ten eerste worden door beluchting vluchtige stoffen uit het water verwijderd, ten tweede wordt het zuurstofgehalte in het veelal zuurstofarme grondwater verhoogd. Dit leidt tot het (gecontroleerd) neerslaan van ijzeroxides, en stimuleert de aanwezigheid van zuurstof tal van biologische processen wat kan leiden tot omzetting of afbraak van stoffen.
- *pH aanpassing.* De zuurgraad van het water wordt aangepast tot een niveau dat gunstig is voor biologische afbraakprocessen. pH aanpassingen worden ook gedaan in het kader van onthardingsprocessen en/of de om de verzadigingsindex in het juiste gebied brengen.
- *Ontharding.* Door de verwijdering van calcium en magnesium ionen worden afzettingen in leidingen voorkomen.
- *Actief kool.* Verwijderen van opgelost organisch materiaal (zoals humuszuren) en organische en anorganische microverontreinigingen met behulp van sorptie.

Met behulp van deze gegevens zou dus ook kunnen worden onderzocht welke (combinatie van) zuiveringstechnieken het meest effectief zijn. In de huidige studie is dit niet gebeurd.

3.3.4 Subscore TotalAntropogenic

Het bepalen of features van antropogene oorsprong zijn is complex. Het kan echter relevant zijn om te weten of features verdacht van antropogene oorsprong zijn omdat dit 1) iets vertelt over de kwetsbaarheid van een winning of een gedeelte van een wingebed voor antropogene invloeden, 2) mogelijkheid biedt om bronmaatregelen in de regio van de onderzochte winningen te nemen, of 3) onderzoek te doen naar antropogene stoffen en hun oorsprong. In de winningen worden verschillende typen zuiveringen en productieprocessen toegepast waarbij verschillende materialen en systemen worden gebruikt, daarbij kunnen stoffen worden omgezet of toegevoegd. De inschatting of een stof (mogelijk) van antropogene oorsprong is op drie manieren bepaald.

A) Correlatie tussen onbekende features en bekende antropogene stoffen (bestrijdingsmiddelen, geneesmiddelen, industriële stoffen). In paragraaf 2.5 wordt uitgelegd hoe de correlatie tussen bekende stoffen en onbekende stoffen (features) is gedaan. Als de correlatiecoëfficiënt tussen het voorkomen van een feature en één of meerdere antropogene stoffen groter of gelijk aan 0,6 is wordt dit als een positieve correlatie gezien en krijgt deze stof de score 1, als niet aan dit criterium wordt voldaan krijgt de stof de score 0. In de positieve modus is de correlatie bij 25 % van de features groter of gelijk aan 0,6. In de negatieve modus is dit zelfs 55%. Dit criterium belicht dus een relevant deel van de aangetroffen features.

B) Inschatting antropogene herkomst op basis van kennis en karakteristieken van de bron. In paragraaf 0 wordt uitgelegd hoe de correlatie tussen bekende stoffen en onbekende stoffen (features) is gedaan. Op basis van deze beoordeling zijn de features ingedeeld in de volgende klassen, en zijn voor deze klassen weegfactoren gekozen:

- Mogelijk antropogeen (weging 1)
- Misschien antropogeen (weging 0,5)
- Mogelijk natuurlijk (weging 0,2)
- Waarschijnlijk natuurlijk (weging 0,1)
- Geen match (worst case; weging 1)

Doordat 84,2% (positieve modus) en 22,6% (negatieve modus) van de features geen match opleverden en een groot deel (resp. 14,3% en 69,1 %) van de resterende features als mogelijk antropogeen aangemerkt. Daardoor is deze factor nog beperkt onderscheidend. In totaal kreeg 98,5% (positieve modus) en kreeg 91,7% (negatieve modus) namelijk de score 1. Dit is voor een deel te wijten aan de hoge fractie die geen match op leverde (positieve modus) maar ook omdat de gekozen criteria in deze dataset een beperkt onderscheid maken (zowel positieve als negatieve modus).

C) Bevat de stof een halogeen (Cl, Br). In paragraaf 2.5 wordt uitgelegd hoe de correlatie tussen bekende stoffen en onbekende stoffen (features) is gedaan. Stoffen met halogenen zijn vaak van antropogene oorsprong of ontstaan door antropogene processen omdat het maken van gehalogeneerde verbindingen veel energie kost. De massaspectra van moederstof en fragmenten kan door de typerende isotopenpatronen van halogenen met een zeer specifiek massaverschil worden herkend. Stoffen met een halogeen krijgen een score van 1, stoffen zonder halogeen een score van 0. In de positieve modes bevatte 2,3% van de features bevatte een halogeen en in de negatieve modus 3,1%. Dit criterium belicht dus slechts een beperkte fractie van alle features.

Vervolgens worden deze verschillende benaderingen als volgt samengevoegd om tot één score voor de (kans op) antropogene herkomst te komen. Deze score is als volgt berekend:

$$TotalAntropogenic (maximum is 1) = 0,5 * factor A + 0,5 * factor B + factor C$$

Waarbij de score maximaal de waarde 1 krijgt, dus als factor C "1" is tellen de andere factoren niet meer mee.

3.3.5 Uitsluitingscriterium TotalRepeat

Één enkele waarneming kan het gevolg zijn van een verontreiniging bij bemonstering, vervoer, opwerking, HRMS of data analyse, daarom is het vaker aantreffen van een feature van belang voor de betrouwbaarheid dat deze feature daadwerkelijk in het water aanwezig was. In deze studie zijn daarom alle features die maar één of twee keer zijn aangetroffen niet meegenomen in verdere analyse.

3.3.6 Uitsluitingscriterium Nolsotope

Om te voorkomen dat isotopen (zoals bijvoorbeeld C₁₂ en C₁₃, CL₃₅ en CL₃₇ etc.) als afzonderlijke features meedoen in de prioritering zijn de isotopen bepaald en uit de totaalijst geschrapt. Alle isotopen (score 0) anders dan de mono-isotopische piek (score 1) zijn niet meegenomen.

3.4 Prioriteren van Toxische eigenschappen

Idealiter zou ook een toxicologische voorselectie wenselijk zijn. Daartoe is geëvalueerd of er een correlatie is tussen de toxiciteit van een stof en eigenschappen die met behulp van hoge resolutie massa spectrometrie kunnen worden bepaald zonder identificatie, i.e. accurate massa en brutoformule. Omdat er geen correlatie is gevonden wordt dit aspect echter niet in de prioritering meegenomen (beschreven hieronder).

3.4.1 Hypothese: Indicatieve toxiciteit gebaseerd op brutoformule

De mogelijke potentie van een stof om bij blootstelling een negatieve impact op de gezondheid te veroorzaken is een belangrijke factor in het prioriteren van stoffen voor grondige gezondheidskundige risicobeoordeling.

In deze studie is een 'quick and dirty' methode verkend waarin de mogelijk in Vitens bronnen aanwezige stoffen zijn verdeeld over verschillende toxiciteitsklassen. De voorgestelde hypothese is dat de bruto formule van chemische stoffen op enige wijze voorspellend is voor toxiciteit (zie tekst box A). Om deze hypothese te toetsen is de CMR classificatie van Vitens doelstoffen (n=341) vergeleken met een voorgestelde classificatie op basis van de chemische bruto formule. CMR classificatie geeft aan of een stof carcinogeen (kankerverwekkend), mutageen (veranderingen in erfelijke eigenschappen inducerend) en/of reproductie-toxisch (schadelijk voor de voortplanting of het nageslacht) is. Stoffen die 1 of meerdere van deze eigenschappen hebben worden CMR stoffen genoemd, en er wordt onderscheid gemaakt tussen verschillende categorieën van bewijs (categorie 1A: bewijs op basis van gegevens voor de mens; categorie 1B: bewijs op basis van voldoende diergegevens; en categorie 2: bewijs op basis van beperkte diergegevens). In deze analyse zijn ook alle stoffen uit de Classification & Labeling (C&L) inventaris van het European Chemicals Agency (ECHA) meegenomen. In dit verkennend onderzoek zijn verschillende opties voor het analyseren en visualiseren van de verzamelde data toegepast.

Tekst box A. Hypothese Toxiciteitsklassen op basis van bruto formule

Stoffen worden geclassificeerd op basis van het aantal dubbele binding equivalenten (DBE) en de aanwezigheid van bepaalde elementen in de bruto formule.

DBE berekening uit bruto formule

$$DBE = A - (B/2) + (C/2) + 1$$

A = aantal C- of Si-atomen

B = aantal H- of halogeen-atomen

C = aantal N- of P-atomen

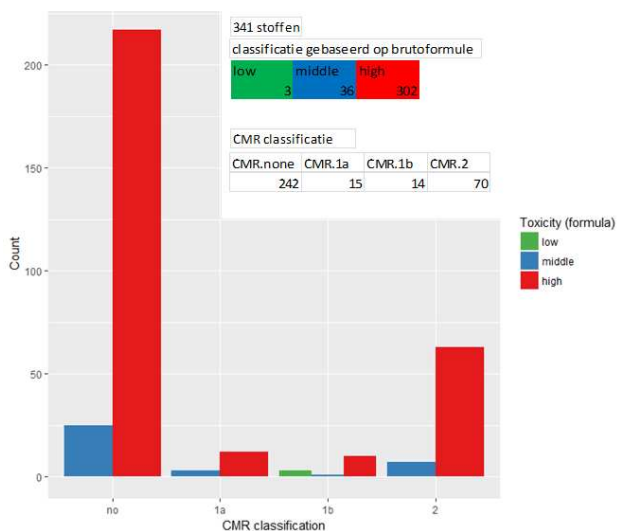
Classificatiehypothese:

'Hoge toxiciteit': DBE>3 + bruto formule bevat C, H, O,N, P en/of halogeen

'Lage toxiciteit': DBE<1 + bruto formule is combinatie van C, H, en/of O

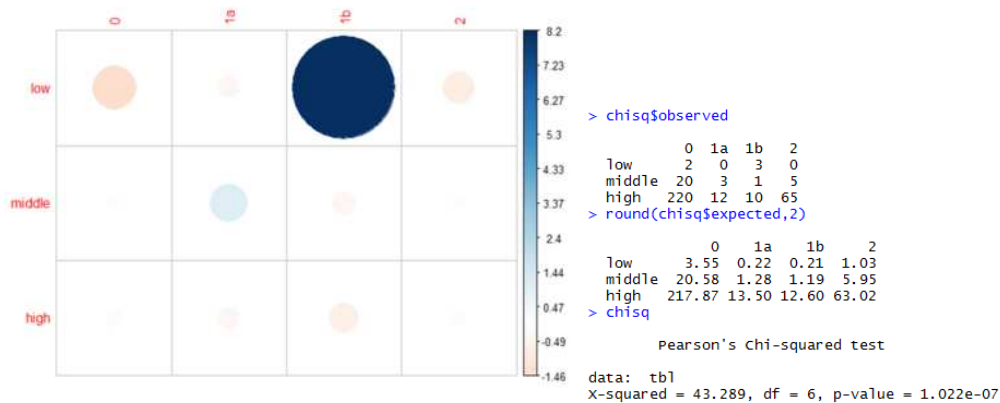
'Middelmatige toxiciteit': rest

Van de 341 Vitens doelstoffen vallen er respectievelijk 3, 36 en 302 in klassen 'Lage tox verwacht', 'Middel tox verwacht' en 'Hoge tox verwacht'. In een vergelijking van CMR classificaties en op bruto formule gebaseerde toxiciteitsklassen (Figuur 3.1) kan echter geen relatie tussen deze factoren worden onderscheiden. Meer dan 200 stoffen zonder CMR classificatie vallen binnen klasse 'Hoge toxiciteit' op basis van bruto formule.



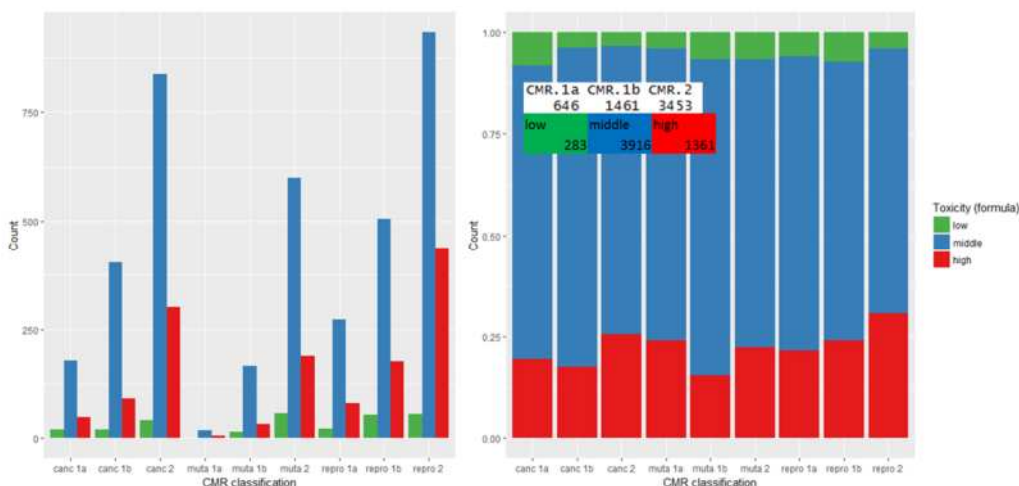
FIGUUR 3-1. VERGELIJKING CMR CLASSIFICATIES EN OP BRUTO FORMULE GEBASEERDE TOXICITEITSKLASSEN VAN DE 341 IN VITENS BRONNEN AANWEZIGE BEKENDE STOFFEN.

De statistische Pearson's Chi-square test kan worden gebruikt om te toetsen of mogelijke verschillen tussen sets niet aan elkaar gerelateerde categorische data toeval zijn. Deze test is gebruikt om te toetsen of CMR classificatie onafhankelijk is van Toxiciteitsklassen op basis van bruto formule (de nulhypothese). Deze nulhypothese moet worden verworpen op basis van de uitkomst van de toets ($p=1.022 \times 10^{-7} < 0.05$), maar de gevonden correlaties zijn niet plausibel; lage toxiciteitsklas correleert met CMR classificatie 1b (Figuur 3.2).



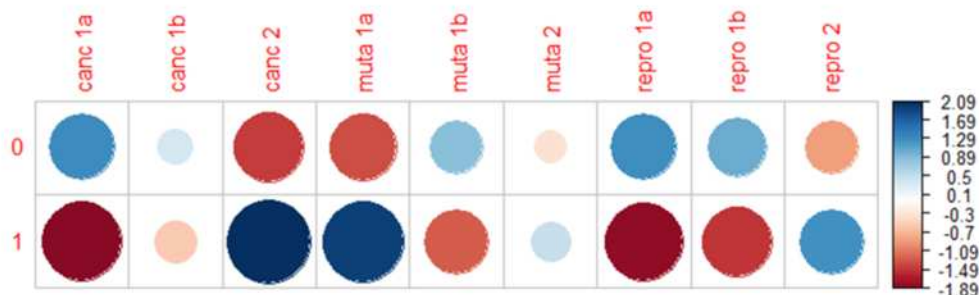
FIGUUR 3-2. CORRELATIES TUSSEN CMR CLASSIFICATIES EN OP BRUTO FORMULE GEBASEERDE TOXICITEITSKLASSEN VAN DE 341 IN VITENS BRONNEN AANWEZIGE STOFFEN. DE GROOTTE VAN DE CIRKEL WEERSPIEGELT DE RELATIEVE BIJDRAGE. DE KLEUR VAN DE CIRKEL GEEFT AAN OF HET EEN POSITIEVE (BLAUWTINTEN) DANWEL NEGATIEVE (ROODTINTEN) ASSOCIATIE BETREFT, WAARBIJ DE GROOTTE VAN DEZE ASSOCIATIE OVEREENKOMT MET DE INTENSITEIT VAN DE KLEUR.

Omdat het aantal doelstoffen ($n=341$) mogelijk te beperkt is om de verwachte correlaties te onderscheiden is een vergelijkbare exercitie uitgevoerd met alle CMR geclassificeerde stoffen uit het C&L inventaris van ECHA (<https://echa.europa.eu/nl/regulations/clp/cl-inventory>, ECHA, 2017) door deze te linken aan bruto formules beschikbaar in de CompTox database van US-EPA (<https://comptox.epa.gov/dashboard>). Voor 5560 stoffen is een mogelijke correlatie onderzocht tussen CMR classificatie (uitgesplitst in classificatie voor carcinogeniteit, mutageniteit en reproductietoxiciteit met verschillende bewijslast) en op bruto formule gebaseerde toxiciteitsklassen (Figuur 3.3). De stoffen met CMR classificatie zijn afkomstig uit de C&L Inventarisatie Database van ECHA (<https://echa.europa.eu/nl/regulations/clp/cl-inventory>). Ook voor deze set stoffen is de Pearson's Chi-square test gebruikt om te toetsen of er een verband is tussen CMR classificatie en Toxiciteitsklassen op basis van bruto formule, waarbij de gevonden correlaties niet plausibel zijn (zie Bijlage I) ondanks een significante uitkomst ($P < 0.05$) van de toets.



FIGUUR 3-3. VERGELIJKING VAN EXACTE (LINKS) EN RELatieve (RECHTS) AANTALLEN STOFFEN MET VERSCHILLENDE CMR CLASSIFICATIE EN BEWIJSLAST EN OP BRUTO FORMULE GEBASEERDE TOXICITEITSKLASSEN VAN CMR GECLASSIFICEERDE STOFFEN (ECHA).

In de prioriteringstool wordt de aanwezigheid van halogenen in de bruto formule meegenomen in de subscore voor de herkomst van de stof (Subscore Antropogene oorsprong), omdat bekend is dat in veel antropogene stoffen halogenen zijn opgenomen voor de stabiliteit van de moleculen. Er is ook verkend of een correlatie bestaat tussen de aanwezigheid van halogenen en CMR classificatie (Figuur 3.4). De gevonden correlaties zijn niet plausibel ondanks een significante uitkomst ($P < 0.05$) van de toets, bijvoorbeeld is er een sterke correlatie tussen CMR classificatie kankerverwekkend 1a en geen halogenen.



FIGUUR 3-4. CORRELATIES TUSSEN AANWEZIGHEID VAN HALOGENEN (0 GEEN HALOGEN, 1 HALOGEN) EN CMR CLASSIFICATIES (ECHA). DE GROOTTE VAN DE CIRKEL WEERSPIEGELT DE RELatieve BIJDRAGE. DE KLEUR VAN DE CIRKEL GEEFT AAN OF HET EEN POSITIEVE (BLAUWTINTEN) DANWEL NEGATIEVE (ROODTINTEN) ASSOCIATIE BETREFT, WAARBIJ DE GROOTTE VAN DEZE ASSOCIATIE OVEREENKOMT MET DE INTENSITEIT VAN DE KLEUR.

Voor de brutoformule $C_6H_{14}O_3$, een voorbeeld bruto formule met classificatie 'Lage tox' volgens de hypothese met betrekking tot toxiciteitsklassen op basis van bruto formule (zie tekst box A), zijn bijpassende stoffen opgezocht in Chemspider (www.chemspider.com). Er passen 136 verschillende stoffen, i.e. structuurformules, bij deze bruto formule. Van de 4 meest geciteerde stoffen uit de 136 vallen 2 in de CRM categorie reprotoxisch 1b. De anderen 2 zijn niet geclassificeerd, omdat bij de ene test data ontbreekt, de andere is wel getest en niet mutageen of reprotoxisch (kankerverwekkend data ontbreekt). Deze uitkomst is in overeenstemming met bovenstaande resultaten die laten zien dat er geen correlaties zijn tussen de toxiciteitsklassen op basis van bruto formule en CMR classificatie van stoffen. De hypothese dat toxiciteit voorspeld kan worden op basis van bruto formule is daarmee weerlegd. Daarom kan in de prioritering op voorhand geen weging worden opgenomen voor toxicologische eigenschappen van features op basis van hun accurate massa, retentietijd, MS 2 spectra en daaruit afgeleide (mogelijke) bruto formule.

3.4.2 Discussie toxicologische eigenschappen

De verschillende CMR classificaties onderscheiden zich op de mate van bewijslast:

- categorie 1A: bewijs op basis van gegevens voor de mens;
- categorie 1B: bewijs op basis van voldoende diergegevens; en
- categorie 2: bewijs op basis van beperkte diergegevens).

Het is echter ook belangrijk te realiseren dat hierin data-beschikbaarheid een belangrijke rol kan spelen. Een stof geclassificeerd als categorie 1B (bewijs op basis van voldoende diergegevens) kan dezelfde mate van carcinogeniteit, mutageniteit of reproductietoxiciteit vertonen ondanks het ontbreken van gegevens voor de mens. Daarnaast zijn bij het onderzoeken van mogelijke correlaties geen stoffen meegenomen zonder CMR classificatie.

De hypothese dat toxiciteit kan worden voorspeld op basis van de bruto formule kan daarbij niet worden bevestigd. Voor risicobeoordeling is het daarom noodzakelijk om de gevonden bruto formules te linken aan een structuurformule van een suspect of bevestigde stof. Daarvan kan de toxiciteit worden opgezocht in databases of gemodelleerd/voorspeld aan de hand van QSAR tools. Voorbeelden hiervan zijn het verzamelen van toxiciteitsinformatie over stoffen uit de ToxCast database met *in vitro* toxiciteitsdata voor grote hoeveelheden stoffen (US EPA, Brunner et al. J HAZMAT, in review), en het schatten van toxiciteit van stoffen met de ToxTree tool.

3.5 Totaalscore prioritering

Na een voorselectie van een minimaal aantal waarnemingen of wel de herhaalbaarheid (TotalRepeat) en het uitsluiten van isotopen (NoIsotope) is de totale prioriteitscore 'Grand Total' is gebaseerd op subscores voor respons (Total Intensity), waarnemingen (TotalFrequency), Significante af- of toename (TotalChange), c) Antropogene oorsprong (TotalAntropogenic) om tot de totale score te komen:

$$Grand\ Total =$$

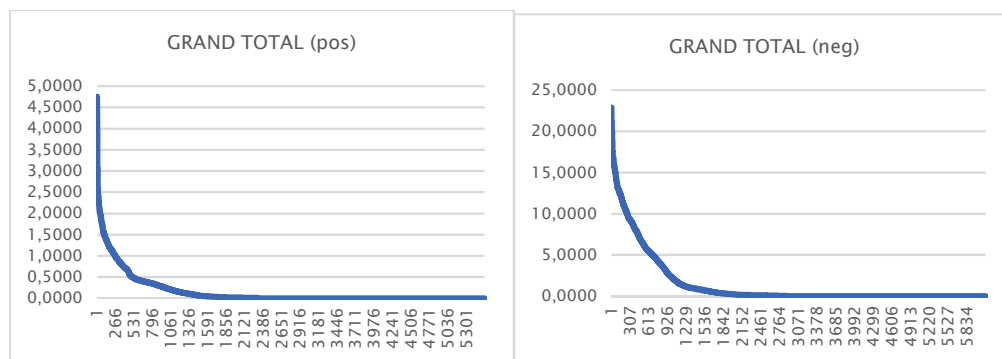
$$TotalRepeat * NoIsotope * (TotalIntensity + TotalFrequency) * TotalChange * TotalAntropogenic$$

Daarbij spelen de intensiteit en voorkomen een belangrijke rol spelen in de prioritering van de feature, omdat hun waarden een breed spectrum bestrijken en worden opgeteld, terwijl de factoren TotalChange en TotalAntropogeen deze waarde wegen door middel van een vermenigvuldiging op basis van eventuele toename of vorming tijdens zuivering en (mogelijke) antropogene oorsprong. TotalRepeat en Nolsotope worden slechts als uitsluitingscriteria gebruikt. Daarbij vallen features af die weinig zijn waargenomen of dat de monoisotopische massa van de feature al elders in de lijst is opgenomen zodat dubbelingen worden voorkomen.

Voor adducten is bewust niet gekozen om deze uit te sluiten in de prioriteringslijst omdat het onderscheid tussen een adduct gevormd tijdens de ionisatie en een unieke feature met een zelfde kenmerkend massa verschil niet *a priori* kunnen worden onderscheiden.

3.6 Geprioriteerde features

Figuur 3-5 laat de distributie van GrandTotal scores voor positieve en negatieve prioritering zien voor respectievelijk 5558 en 6140 features voor de positieve en negatieve ionisatiemodus, waarvan respectievelijk 2301 (41%) en 2886 (47%) een score groter dan 0 hebben. De verdeling laat zien dat er een beperkt aantal stoffen met een hoge score is en veel meer stoffen met een lagere score. De prioriteringsystematiek met weging van verschillende factoren heeft een onderscheidend vermogen tussen de verschillende features. De schaal van de positieve en negatieve ionisatie modus zijn genormaliseerd naar de interne standaard respons. Neburon is in beide modi als interne standaard gebruikt.



FIGUUR 3-5: VERDELING VAN DE TOTALE SCORES VAN DE PRIORITERING IN POSITIEVE EN NEGATIEVE IONISATIEMODUS

Tabel 3.1 en 3.2 geven de features met hoge prioriteit weer in de positieve en negatieve modus. De features zijn geselecteerd op basis van hun score en “expert knowledge” over de mogelijkheid tot het identificeren van deze stoffen. Deze methode verschilt van de methode die wordt gehanteerd in de gerelateerde studie [20] omdat in dit stadium de kans dat er van de suspects een CAS nummer én toxicologische informatie beschikbaar is beperkt wordt geacht. Dit wordt bevestigd door latere matching met beschikbare CAS nummers en analyse van toxicologische gegevens van suspects (met CAS nummer) van de geprioriteerde features (zie paragraaf 0). Bij het ontbreken van toxicologische gegevens van het merendeel van de suspects kan de feature niet toxicologisch worden beoordeeld.

TABEL 3.1 GEPRIORITEERDE FEATURES POSITIEVE MODUS.

P	Feature Nr	Componentnaam	prioriteringscore
1	1504	196.0965 / 11.83	3,93
2	746	284.0974 / 11.16	3,39
3	607	218.1380 / 10.53	3,22
4	525	186.1477 / 12.30	3,09
5	1286	341.1219 / 18.29	2,42
6	354	469.3140 / 36.98	1,81
7	1303	557.2218 / 21.63	2,42
8	548	198.1340 / 12.97	2,67
9	2554	216.0674 / 10.03	1,94
10	2431	421.0951 / 12.61	1,73
11	1207	445.2680 / 14.50	2,40
12	706	261.1221 / 17.57	2,23

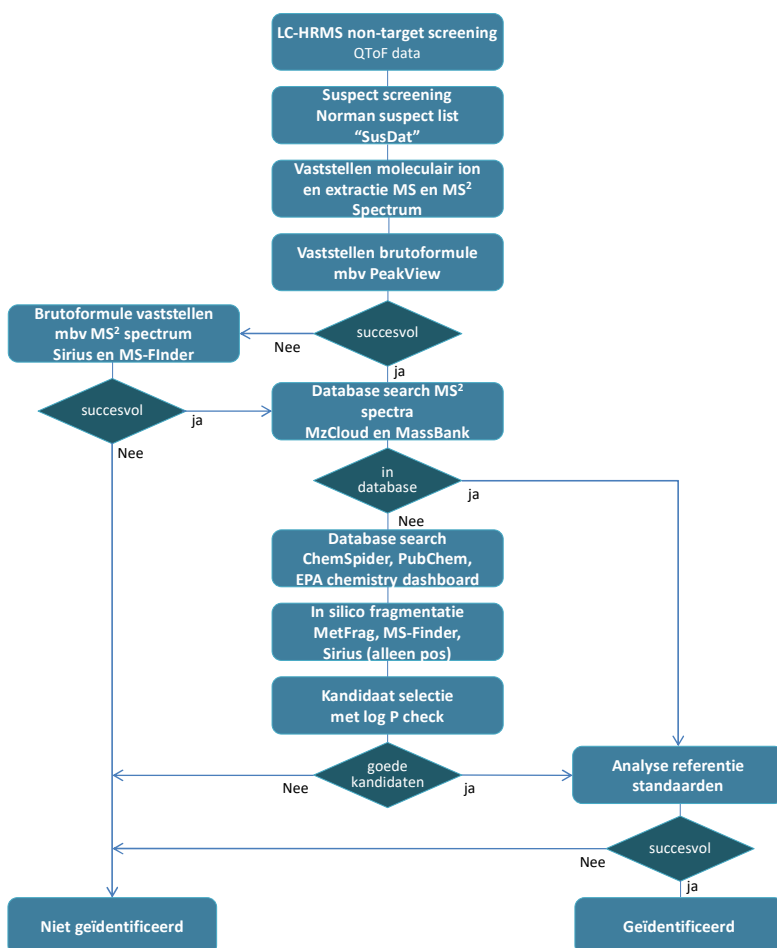
TABEL 3.2 GEPRIORITEERDE FEATURES NEGATIEVE MODUS.

N	Feature Nr	Componentnaam	prioriteringscore
1	1401	285.0802 / 21.16	22,43
2	911	230.0616 / 10.48	21,40
3	1350	273.0089 / 13.93	21,14
4	1442	243.0336 / 12.23	21,08
5	1972	273.0450 / 17.29	20,98
6	475	183.0896 / 13.33	20,09
7	3239	299.0952 / 23.21	19,85
8	1750	271.0653 / 17.51	18,69
9	3266	269.0484 / 24.14	17,48
10	1948	253.0552 / 21.78	17,48

4 Identificatie en evaluatie geprioriteerde stoffen

4.1 Workflow identificatie geprioriteerde stoffen

Om de identiteit te kunnen achterhalen van de geprioriteerde features / stoffen is een identificatie onderzoek uitgevoerd. Voor de identificatie is een workflow opgezet die afgeleid is van de identificatie workflow die binnen het BTO project "Tools voor de identificatie van onbekende verbindingen met hoge resolutie massaspectrometrie data" is ontwikkeld. De identificatie workflow is geoptimaliseerd voor het huidige onderzoek en gebruikte massaspectrometer met bijbehorende massaspectrometrische data. De ontwikkelde workflow voor de identificatie van geprioriteerde stoffen staat weergegeven in Figuur 4-1. Hieronder worden de verschillende stappen in de workflow toegelicht.



FIGUUR 4-1. IDENTIFICATIE WORKFLOW VOOR GEPRIORITEERDE STOFFEN.

4.2 Toepassing

De identificatie workflow is toegepast voor non-target screening data opgenomen met een AB Sciex TripleTof 5600+ massaspectrometer (30.000 FWHM) en bevat full-scan MS spectra alsmede data afhankelijke MS² spectra.

Op de geprioriteerde stoffenlijst is eerst een suspect screening uitgevoerd met behulp van de Norman suspect list "SusDat". Hierbij waren enkele mogelijke hits gevonden voor verdere identificatie, welke later in het identificatieproces als mogelijke kandidaten zijn meegenomen. Na de suspect screening werden de brutoformules van de geprioriteerde stoffen vastgesteld. Hiervoor zijn chromatogrammen en massaspectrometrische data geselecteerd waarin de geprioriteerde stoffen in relatief hoge concentraties zijn gedetecteerd, zodoende een informatierijk MS¹ spectrum en MS² spectrum te verkrijgen. Eerst is het MS¹ spectrum bekeken om het moleculaire ion met zekerheid vast te stellen en uit te sluiten dat de gedetecteerde massa geen adduct of fragment ion (in source fragmentatie) is. Voor verdere dataverwerking is vervolgens het MS¹ en MS² spectrum handmatig geëxtraheerd en opgeslagen als peaklist in csv (comma separated value) formaat. De brutoformule is vastgesteld met behulp van PeakView software (AB Sciex), waarin ook een extra adduct-check aanwezig is. Wanneer de brutoformule niet goed kon worden vastgesteld (bijv. meerdere kandidaten of verstoord MS¹ spectrum), is er gebruik gemaakt van de software tools Sirius en MS-Finder, waarmee aan de hand van het MS¹ en MS² spectrum (doormiddel van import MS¹ en MS² peaklist) de brutoformules naar waarschijnlijkheid worden gerangschikt. Na het vaststellen van de brutoformule is er vervolgens een MS² spectra database search uitgevoerd in de mzCloud en MassBank database. Wanneer een verbinding in een van de databases voorkomt kan de identiteit ondubbelzinnig bevestigd worden door de analyse van een referentiestandaard. Indien de verbinding niet aanwezig was in een van de MS² databases werd de zoektocht uitgebreid naar databases zoals ChemSpider, PubChem en EPA chemistry dashboard. Hierbij werd dan gezocht op de vastgestelde brutoformules, waarna de gevonden verbindingen werden gesorteerd op relevantie doormiddel van referenties of PubMed filters toe te passen. Hierdoor werd inzicht verkregen in het aantal mogelijk kandidaten, en of er ook bekende/relevante kandidaten aanwezig waren. Uiteindelijk zijn er veel kandidaat stoffen gevonden per feature en is er gebruik gemaakt van *in silico* fragmentatie om het aantal kandidaat stoffen te reduceren. Hiervoor zijn de *in silico* fragmentatie tools MetFrag, MS-Finder en Sirius (alleen positieve mode) gebruikt. Vervolgens zijn de hoogst gerankte kandidaten tussen de *in silico* fragmentatie tools vergeleken, om tot een goede kandidaatselectie te komen. Daarna zijn de log P waarden van de kandidaten berekend (ACD/Labs) om te bepalen of de retentietijd plausibel is bij de log P waarde van de kandidaatstof(fen). Daarvoor is een correlatie tussen de RT en de log P voor 20 bekende stoffen gemaakt, en is voor de kandidaat stoffen een bandbreedte van 2 log-eenheden om deze correlatie toegestaan. Aan de hand van deze stappen is een kandidatenlijst opgesteld voor de geprioriteerde verbindingen.

4.3 Resultaten identificatie geprioriteerde stoffen

De identificatie van de geprioriteerde stoffen is uitgevoerd volgens de identificatie workflow voor 12 positieve en 10 negatieve geprioriteerde stoffen Tabel 3-1 en 3-2. De resultaten van het identificatieproces staan vermeld in Tabel 4-1 en Tabel 4-2.

TABEL 4-1 RESULTATEN IDENTIFICATIE GEPRIORITEERDE STOFFEN (POSITIEVE ANALYSE).

P	Feature Nr	Componentnaam	Brutoformule	Identiteit	Opmerkingen
1	1504	196,0965 / 11,83	C ₁₀ H ₁₃ NO ₃	Kandidaatstoffen: Metirosine (672-87-7) N-methyltyrosine (537-49-5)	Matig MS ² spectrum, veel hits/kandidaten (> 8000)
2	746	284,0974 / 11,16	C ₉ H ₁₄ O ₉	Onbekend	NH ₄ adduct van m/z 267.0710, geen goede kandidaten <i>in silico</i> fragmentatie
3	607	218,1380 / 10,53	C ₁₀ H ₁₉ NO ₄	Kandidaatstoffen: 2-Aminodecanedioic acid (3721-88-8) N-(tert-butoxycarbonyl)valine (54895-12-4)	Veel kandidaten (> 4000)
4	525	186,1477 / 12,30	C ₁₀ H ₁₉ NO ₂	Kandidaatstof: (1-Isopropyl-4-piperidiny)acetic acid (754183-67-0)	Veel mogelijke kandidaten met <i>in silico</i> fragmentatie (>11000)
5	1286	341,1219 / 18,29	C ₁₅ H ₁₈ N ₄ O ₄ ?	Onbekend	Gedetecteerd op Na adduct! Matig MS ² spectrum (spectrum van Na adduct)
6	354	469,3140 / 36,98	C ₂₂ H ₄₀ N ₆ O ₅	Onbekend	matig MS ² spectrum, geen goede kandidaten (mogelijk Na adduct)
7	1303	557,2218 / 21,63	C ₂₄ H ₃₈ O ₁₃ ?	Onbekend	Is m/z van Na adduct! geen goede hits met <i>in silico</i> fragmentatie (NH ₄ MS ²)
8	548	198,1340 / 12,97	C ₈ H ₁₅ N ₃ O	Hydroxyatrazine	Bevestigd via mzCloud en MassBank MS ² spectrum
9	2554	216,0674 / 10,03	C ₉ H ₁₃ NO ₃ S	Kandidaatstoffen: 3-Anilino-1-propanesulfonic acid (72943-20-5) 3-(Ethylamino)-4-methylbenzenesulfonic acid (98-40-8)	
10	2431	421,0951 / 12,61	C ₁₉ H ₁₈ N ₄ O ₄ S	Onbekend	Is m/z van Na adduct! matig (NH ₄) MS ² spectrum, geen goede kandidaten
11	1207	445,2680 / 14,50	C ₂₅ H ₃₆ N ₂ O ₅	Onbekend	matig MS ² spectrum, geen goede kandidaten
12	706	261,1221 / 17,57	C ₁₄ H ₁₆ N ₂ O ₃	Onbekend	Veel hits in databases (> 20000), hierdoor geen onderscheidende kandidaten

TABEL 4-2 RESULTATEN IDENTIFICATIE GEPRIORITEERDE STOFFEN (NEGATIEVE ANALYSE).

N	Feature Nr	Componentnaam	Brutoformule	Identiteit	Opmerkingen
1	1401	285,0802 / 21,16	C ₁₃ H ₁₈ O ₅ S	Kandidaatstof: Ethofumesate (26225-79-6)	Top kandidaat met Metfrag en MS-Finder, wordt wel voornamelijk in pos mode gemeten
2	911	230,0616 / 10,48	?	Onbekend	Ongevoelig (MS' piek 3500 counts)
3	1350	273,0089 / 13,93	C ₁₀ H ₁₀ O ₇ S	Kandidaatstof: Ferulic acid 4-sulfate (86321-29-1)	Ferulic Acid 3-Sulfate ook mogelijk
4	1442	243,0336 / 12,23	C ₁₀ H ₁₂ O ₅ S	Kandidaatstoffen: Methyl 2-methoxy-5-(methylsulfonyl)benzoate (63484-12-8) 3-[(4-Methoxyphenyl)sulfonyl]propanoic acid (91062-23-6) 2-[4-(Methoxycarbonyl)phenyl]ethanesulfonic acid (756469-30-4)	Beperkte kans op succes
5	1972	273,0450 / 17,29	C ₁₁ H ₁₄ O ₆ S	Kandidaatstof: 3-[(3,4-Dimethoxyphenyl)sulfonyl]propanoic acid (100059-84-5)	Beperkte kans op succes
6	475	183,0896 / 13,33	C ₇ H ₁₂ N ₄ O ₂	Kandidaatstof: metribuzin-diketo	Top kandidaat met Metfrag en MS-Finder.
7	3239	299,0952 / 23,21	C ₁₄ H ₂₀ O ₅ S?	Onbekend	Brutoformule niet geheel zeker, meerdere onbekende kandidaten, niet te koop
8	1750	271,0653 / 17,51	C ₁₂ H ₁₆ O ₅ S	Onbekend	Meerdere kandidaten met goede match, maar allen zeer onwaarschijnlijk of niet te koop.
9	3266	269,0484 / 24,14	C ₁₂ H ₁₄ O ₅ S	Kandidaatstoffen: [(2S)-5-Oxotetrahydro-2-furanyl]methyl 4-methylbenzenesulfonate (58879-34-8) Genistein (446-72-0) Apigenin (520-36-5)	Meerdere kandidaten met goede match, geen onderscheidende kandidaten
10	1948	253,0552 / 21,78	C ₁₂ H ₁₄ O ₄ S	Onbekend	Meerdere matches met MetFrag, maar allen onbekend, en niet verkrijgbaar

De identificatie van de geprioriteerde stoffen heeft geleid tot de ondubbelzinnige identificatie van één verbinding, namelijk hydroxyatrazine. Deze verbinding is bevestigd doormiddel van een zeer goede match van het MS² spectrum met de aanwezige MS² spectra in de mzCloud en Massbank databases. De identiteit van hydroxyatrazine was op voorhand al bekend voor een select aantal personen, maar is uiteindelijk als onbekende aangemerkt om zo de gehele identificatie workflow te toetsen. Verder heeft de identificatie geleid tot mogelijke kandidaat stoffen voor vier verbindingen geanalyseerd met de positieve analyse en voor zes verbindingen geanalyseerd met de negatieve analyse. De identiteiten van de kandidaat stoffen kunnen pas bevestigd of ontkracht worden door middel van het aanschaffen en analyseren van deze referentiestoffen. Zie bijlage II voor een overzicht van de beschikbaarheid en leveranciers van deze kandidaat stoffen. Totdat de referentiestandaarden geanalyseerd zijn, blijft het een voorlopige identificatie waarbij de identiteit van de onbekende stof niet volledig is opgehelderd.

Voor zeven geprioriteerde stoffen gedetecteerd in de positieve ionisatiemodus en vier geprioriteerde stoffen gedetecteerd in de negatieve ionisatie modus konden geen beperkt aantal goede kandidaat stoffen worden gevonden. Dit kan mogelijk verklaard worden door de volgende oorzaken:

- Lage intensiteit voor veel geprioriteerde verbindingen, waardoor er veel achtergrond/verstoringen zijn in het MS¹ en MS² spectrum. Dit kan voorkomen worden door een hogere drempelwaarde voor de intensiteit te hanteren tijdens de prioritering.
- Weinig MS² fragmentionen waardoor er geen goede *in silico* fragmentatie kan worden uitgevoerd.
- Adductvorming, hierdoor is de intensiteit van het molecuul ion en bijbehorende MS² spectrum vaak te laag voor een goede identificatie.
- Beperkte resolutie voor hoge molecuulmassa's. Vanwege het ontbreken van de isotoop fijnstructuur (resolutie > 100.000 FWHM benodigd) is het niet altijd goed mogelijk om brutoformule vast te stellen (in het bijzonder verbindingen met een hoge molecuulmassa).
- Mogelijk zijn relatief veel geprioriteerde stoffen, transformatieproducten of anderszins moeilijk te identificeren stoffen, welke niet aanwezig zijn in databases waardoor er geen referentie informatie is om deze stoffen eenvoudig(er) te identificeren.

4.4 Bevestiging geprioriteerde suspects

Voor de bevestiging van de geprioriteerde stoffen zijn van de mogelijke kandidaat stoffen (zie 4.2) de referentiestandaarden aangeschaft. Niet alle kandidaat stoffen waren verkrijgbaar. De verkregen referentiestandaarden zijn vervolgens onder dezelfde condities als de monsters waarin de geprioriteerde features zijn gedetecteerd geanalyseerd met behulp van de QToF Triple TOF 5600+ massaspectrometer. Wanneer de accurate massa, retentietijd en MS² spectrum van de referentiestandaard overeenkomt met die in het desbetreffende monster, is de stof ondubbelzinnig bevestigd. De resultaten van de bevestiging van de kandidaat stoffen staan vermeld in Tabel 4-3 en 4-4 voor de positieve en negatieve analyse.

TABEL 4-3 RESULTATEN BEVESTIGING KANDIDAATSTOFFEN POSITIEVE ANALYSE

Kandidaatstoffen	Cas nr	Brutoformule	Feature nr	RT monster	RT standaard	Bevestigd	
						RT	MS2
N-Methyltyrosine	537-49-5	C ₁₀ H ₁₃ NO ₃	1504	11,83	5,96	nee	nee
N-(tert-buroxycarbonyl)valine	54895-12-4	C ₁₀ H ₁₉ NO ₄	607	10,53	20,34	nee	nee
3-Anilino-1-propanesulfonic acid	72943-20-5	C ₉ H ₁₃ NO ₃ S	2554	10,03	9,53	nee	nee
3-(Ethylamino)-4-methylbenzenesulfonic acid	98-40-8	C ₉ H ₁₃ NO ₃ S	2554	10,03	9,18	nee	nee
(1-Isopropyl-4-piperidinyl)acetic acid	754183-67-0	C ₁₀ H ₁₉ NO ₂	525	12,3	7,29	nee	nee

TABEL 4-4 RESULTATEN BEVESTIGING KANDIDAATSTOFFEN NEGATIEVE ANALYSE

Kandidaatstoffen	Cas nr	Brutoformule	Feature nr	RT monster	RT standaard	Bevestigd	
						RT	MS2
[(2S)-5-Oxotetrahydro-2-furanyl]methyl 4-methylbenzenesulfonate	58879-34-8	C ₁₂ H ₁₄ O ₅ S	3266	n.a.w.	24,14	nee	nee
Genistein	446-72-0	C ₁₅ H ₁₀ O ₅	3266	20,38	24,14	nee	nee
Apigenin	520-36-5	C ₁₅ H ₁₀ O ₅	3266	20,27	24,14	nee	nee
Ethofumesate	26225-79-6	C ₁₃ H ₁₈ O ₅ S	1401	n.a.	21,16	nee	nee
Methyl 2-methoxy-5-(methylsulfonyl)benzoate	63484-12-8	C ₁₀ H ₁₂ O ₅ S	1442	n.a.	12,23	nee	nee
Ferulic acid 4-sulfate dinatrium	86321-29-1	C ₁₀ H ₁₀ O ₇ S	1972	13,53	17,29	nee	nee
Metribuzin-diketo	56507-37-0	C ₇ H ₁₂ N ₄ O ₂	475	16,43	13,33	nee	nee

n.a. = niet aanwezig

De identificatie van deze stoffen was hiermee niet succesvol, voor geen van de geprioriteerde stoffen kon doormiddel van kandidaat stoffen de identiteit bevestigd worden. Voor alle kandidaat stoffen kwam het MS² spectrum niet overeen met die van de geprioriteerde stof. Op voorhand was ingeschat dat enkele verbindingen doormiddel van deze aanpak geïdentificeerd zouden kunnen worden op basis eerdere studies [21]. Wat waarschijnlijk bijdraagt aan de lagere succes rate van de identificatie, is dat op voorhand de "eenvoudig" te identificeren stoffen al zijn geïdentificeerd en om die reden uit de lijst met geprioriteerde stoffen zijn gehaald. Daardoor bleven de relatief lastig te identificeren stoffen over. Een andere mogelijke oorzaak is dat er tijdens het relatief lange verblijf van water in de ondergrond relatief veel transformatie producten zijn ontstaan en geprioriteerd. Deze stoffen staan vaak niet in databases en ook hun referentie standaarden zijn vaak niet beschikbaar. Hierdoor kunnen deze stoffen niet goed geïdentificeerd worden. Verder waren de intensiteiten van veel geprioriteerde stoffen vrij laag, waardoor er niet altijd een goed MS¹ spectrum (vrij van storingen) en goed MS² spectrum met voldoende fragmenten (fragmenten > 6) waren opgenomen. Deze oorzaken kunnen grotendeels al verholpen worden door een hogere drempelwaarde voor de intensiteit te hanteren tijdens de prioritering.

Een mogelijke verbeterpunt voor een volgend onderzoek is om de herziende workflow te gebruiken (zie 4.6), die meer nadruk legt op het minimale aantal MS² fragmentionen en eventuele heranalyse van de monsters waarin de geprioriteerde stof is aangetroffen. Vaak kan er veel winst worden behaald doormiddel van een heranalyse, door gerichte MS¹ en MS² acquisitie uit te voeren, waardoor gevoeligere en rijkere MS¹ en MS² worden verkregen.

4.5 Mogelijkheden voor verdere identificatie

Voor het overgrote deel van de (nog) niet geïdentificeerde stoffen zijn er voldoende aanknopingspunten voor verdere identificatie. In Tabel 4-5 en 4-6 staan mogelijke oorzaken waarom een verbinding niet geïdentificeerd is, en welke vervolgstappen zinvol zijn om tot betere kandidaten, minder kandidaten of zelfs de identiteit van de stof te komen.

TABEL 4-5 VERBETERINGSPUNTEN VOOR VERDERE IDENTIFICATIE (POSITIEVE ANALYSE).

P	Feature Nr	Aantal MS ² fragmenten	Gevoeligheid MS ²	Oorzaak	Oplossing
1	1504	2	matig	Weinig fragmenten en gevoeligheid	Analyse met hogere CE en/of concentratie
2	746	8	slecht	Adductvorming en gevoeligheid MS ² spectrum (NH ₄ adduct)	Heranalyse met ander eluens en/of concentratie
3	607	7	redelijk	Geen, mogelijk transformatie product	
4	525	8	redelijk	Geen, mogelijk transformatie product	
5	1286	2	matig	Adductvorming, weinig fragmenten (Na adduct) en matig MS ¹ spectrum	Heranalyse met ander eluens en/of hogere resolutie
6	354	3	slecht	Weinig fragmenten en gevoeligheid, mogelijk adduct van Na	Heranalyse met ander eluens, concentratie, hogere CE
7	1303	6	redelijk	Adductvorming, matig MS ¹ spectrum	Heranalyse met ander eluens en/of hogere resolutie
8	548	5	redelijk	nvt	nvt
9	2554	4	redelijk/goed	beperkt aantal fragmenten	Analyse met hogere CE
10	2431	10	matig/slecht	Adductvorming, matig MS ¹ spectrum, veel fragmenten (NH ₄ adduct) net boven ruis niveau	Heranalyse met ander eluens en/of hogere resolutie
11	1207	2	redelijk	Weinig fragmenten	Analyse met hogere CE
12	706	6	goed	Te veel kandidaten, meer fragmenten nodig	Analyse met hogere CE

TABEL 4-6 VERBETERINGSPUNTEN VOOR VERDERE IDENTIFICATIE (NEGATIEVE ANALYSE).

N	Feature Nr	Aantal MS ² fragmenten	Gevoeligheid MS ²	Oorzaak	Oplossing
1	1401	8	matig	Gevoeligheid	Concentratie
2	911	1	slecht	Gevoeligheid	Concentratie
3	1350	8	redelijk	Geen, mogelijk transformatieproduct	
4	1442	4	matig	Gevoeligheid, beperkt aantal fragmenten	Analyse met hogere CE en/of concentratie
5	1972	10	redelijk/goed	Geen, mogelijk transformatieproduct	
6	475	4	redelijk	beperkt aantal fragmenten	Analyse met hogere CE
7	3239	4	goed	beperkt aantal fragmenten, matig MS ¹ spectrum	Analyse met hogere CE en/of hogere resolutie
8	1750	6	redelijk	Geen, mogelijk transformatieproduct	
9	3266	8	goed	Matig, MS ¹ spectrum	Analyse met hogere resolutie
10	1948	9	goed	Geen, mogelijk transformatieproduct	

In bijna alle gevallen leidt een hogere concentratie van een geprioriteerde stof tot een verbeterde identificatie. De brutoformule is dan eenvoudig vast te stellen aan de hand van het MS¹ spectrum, en het MS² spectrum verbetert dan ook aanzienlijk. Er zijn dan meer fragmentionen zichtbaar maar ook de kans op een “vervuild” MS² spectrum van een mogelijke stoorverbinding wordt kleiner. De eenvoudigste manier om de concentratie te verhogen is doormiddel van vaste fase extractie. Of dit tot de mogelijkheden behoort, is afhankelijk is van de polariteit van de verbinding en voor welke vaste fase (extractiemateriaal) gekozen wordt. Hierbij kan als criterium de zogenoemde *k* factor gebruiken. Dit is de verhouding tussen retentietijd van de injectiepiek en de betreffende stof. Met een *k* factor van twee is het aannemelijk dat een solid phase materiaal, die werkt doormiddel van hydrofobe interacties de stof kan extraheren.

$$k = \frac{RT_{stof} - RT_{injectiepiek}}{RT_{injectiepiek}}$$

De identificatie kan verder verbeterd worden door een andere collision energie voor een onbekende verbinding te kiezen. Vaak is dit een hogere collision energie waardoor er meer fragmentionen in het MS² spectrum aanwezig zijn en kunnen worden gemeten. Dit helpt vooral bij het vergelijken van de resultaten met *in silico* fragmentatie doordat de mogelijke kandidaten beter “gematcht” kunnen worden en er uiteindelijk minder mogelijke kandidaat stoffen overblijven. Verder kan de gevoeligheid van het MS² spectrum van een onbekende verbinding verbeterd worden door een extra analyse uit te voeren waarin het MS² spectrum continu opgenomen wordt in combinatie met lange accumulatie tijden. Bij de non-target analyses is de apparatuur ingesteld om een breed spectrum aan massa's (M/Z) te analyseren. Bij her-analyse kan gericht op een onbekende stof (met een massa) worden gescreend, waardoor de gegevensdichtheid sterk verbetert. Hierdoor wordt de signaal/ruis verhouding in het MS² spectrum substantieel beter.

Vooral in de positieve analyse is er relatief veel adductvorming geconstateerd van ammonium (NH₄) en Natrium (Na) adducten. Adductvorming kan de identificatie bemoeilijken doordat maar een (klein) deel van de stof als moleculair ion wordt gemeten (M+H⁺ of M-H). Het gevolg hiervan kan zijn dat er niet altijd een MS² spectrum wordt opgenomen van het moleculaire ion of dat deze een veel lagere intensiteit heeft. Hierdoor kan de identificatie soms alleen worden uitgevoerd met het MS² spectrum van het ammonium of natrium adduct te gebruiken. Ammonium adducten hebben meestal nog een bruikbaar MS² spectrum voor identificatie doeleinden, maar spectra van Na adducten zijn doorgaans niet bruikbaar. Dit kan mogelijk opgelost worden door een her-analyse onder andere omstandigheden (ionisatie), en samenstelling van eluens (minder natrium of ammonium) waardoor de adductvorming wordt beperkt, en de respons en spectra van het moleculaire ion verbeterd. Mocht dit niet werken kan er gekozen worden voor een andere eluens samenstelling om de adductformatie te beperken. Denk hierbij het gebruik van andere buffers, zuren en oplosmiddelen.

Bij een aantal onbekende verbindingen waren ook verstoorde MS¹ spectra geconstateerd, waarbij er stoorpieken/achtergrond aanwezig was op de M+1 en M+2 isotoop posities. Dit bemoeilijkt de identificatie, vooral verbindingen met relatief hoge molecuulmassa's. Een oplossing hiervoor is een extra analyse uit te voeren bij een aanzienlijk hogere resolutie (> 100.000) met een andere massaspectrometer, waardoor de verschillende stoormassa's beter van elkaar gescheiden kunnen worden.

Voor een aantal onbekende verbindingen waren er geen aanwijsbare oorzaken waardoor een identificatie niet goed kon worden uitgevoerd. Deze verbindingen hadden allen veel MS² fragmentionen (≥ 6) maar dit leidde ondanks het gebruik van *in silico* fragmentatie tools

niet tot goede kandidaten. Een mogelijke verklaring hiervoor is dat het in deze gevallen om transformatieproducten gaat. Transformatieproducten staan vaak (nog) niet vermeld in databases of er zijn nog geen referentiestandaarden beschikbaar, waardoor ze niet geïdentificeerd kunnen worden. Voor de geprioriteerde stoffen is dit onderzocht door te zoeken naar bekende massaverschillen van transformatieproducten waardoor potentiële moederstoffen binnen de dataset gevonden kunnen worden. De resultaten hiervan staan vermeld in bijlage III.

4.6 Analyse van geprioriteerde stoffen met de Orbitrap Fusion massaspectrometer

In de voorgaande paragraaf (4.4) is vastgesteld dat er voor de meerderheid van de niet geïdentificeerde geprioriteerde stoffen genoeg aanknopingspunten zijn voor een verdere identificatie. Een van deze aanknopingspunten is de analyse van de geprioriteerde stoffen op zeer hoge resolutie (>100.000 FWHM). Om dit te testen is besloten om een aantal geprioriteerde stoffen te selecteren die zeer waarschijnlijk baat hebben bij een analyse op een hogere resolutie en deze dan te analyseren met een Orbitrap Fusion massaspectrometer. Met de Orbitrap Fusion zijn resoluties mogelijk van 500.000 FWHM (bij m/z 200). Hiervoor zijn vier geprioriteerde stoffen geselecteerd (zie Tabel 4-7).

TABEL 4-7 GESELECTEERDE STOFFEN VOOR ANALYSE MET DE ORBITRAP FUSION.

Nr	Feature nr.	Ionisatie mode	Locatie	Type monster	Datum monsternamen
P7	1303	positief	Pb. Wierden	Winput 1975-19	06-03-2018
P10	2431	positief	Pb. Bremerberg	Reinwater	05-03-2018
N7	3239	negatief	Pb. Laren	Reinwater	05-03-2018
N9	3266	negatief	Pb. Laren	Ruw filter 2	07-03-2018

De stoffen met feature nr. 1303 (P7) en 2431 (P10) die gedetecteerd zijn in de positieve mode zijn geselecteerd omdat hiervan alleen een matig fullscan (MS¹) spectrum beschikbaar is. Dit komt doordat tijdens de QTOF analyse er veel hinder werd ondervonden van adductvorming ([M+Na]⁺ en [M+NH₄]⁺) voor deze twee stoffen. Hierdoor was het geprotoneerd molecuul ion zeer laag in intensiteit in vergelijking met de adducten, waardoor het vaststellen van de brutoformule werd bemoeilijkt en er geen goede MS² spectra beschikbaar waren van het geprotoneerd molecuul ion. De stoffen met feature nr. 3239 (N7) en 3266 (N9) die gedetecteerd zijn in de negatieve mode zijn gekozen omdat er veel achtergrond aanwezig was in het MS¹ spectrum, waardoor niet met zekerheid kon worden vastgesteld of er ook een zwavel isotoop aanwezig was.

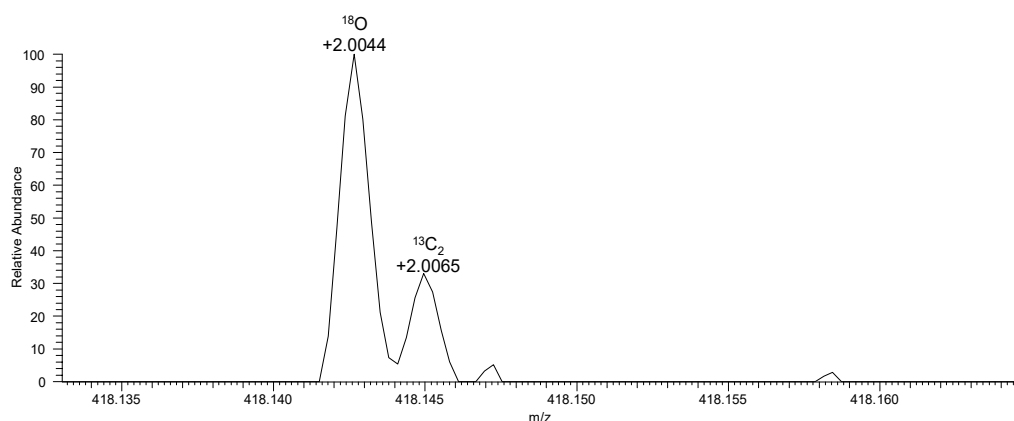
Voor de analyse met de Orbitrap Fusion zijn de locaties opnieuw bemonsterd. Om de kans op een succesvolle identificatie te vergroten, is er voor gekozen om de monsters met directe injectie te analyseren maar ook te concentreren met vaste fase extractie (SPE) en het bijbehorende extract te analyseren. De vaste fase extractie is uitgevoerd met een Oasis HLB vaste fase kolom, waarmee 500 mL monster is geconcentreerd tot 1 mL (zie bijlage IV). De directe injectie analyse en analyse van het extract zijn uitgevoerd volgens de non-target screening methode van KWR (zie bijlage IV). De resultaten van de Orbitrap Fusion analyse van de vier monsters staan hieronder per stof (feature) beschreven.

Feature 1303 (P7)

Met de QToF analyse was feature 1303 gedetecteerd op het Na adduct met een m/z 557,2218, waarbij het geprotoneerd molecuul ion (m/z 535,2372) een zeer lage intensiteit had van circa 2% ten opzichte van het Na adduct. Tijdens de analyse met de Fusion kon deze stof niet terug worden gevonden in monster Pb. Wierden (Winput 1975-19) met zowel de directe injectie als analyse van het SPE extract. Hierbij is gezocht op de accurate massa's van het geprotoneerd molecuul ion, Na en NH_4 adducten. Mogelijke verklaringen hiervoor zijn de dat de stof niet meer aanwezig is in het bronwater, of dat de gevoeligheid met directe injectie op de Fusion te laag is voor deze stof (100 μL injectie vs 1 mL injectie van de QToF), en dat de verbinding niet geïsoleerd wordt met vaste fase extractie ondank de goede retentie op een C18 kolom (21.63 minuten).

Feature 2431 (P10)

De stof met feature nr. 2431 was met de QToF analyse ook op het Na adduct gedetecteerd, waarbij het geprotoneerd molecuul ion een lagere intensiteit heeft van circa 10% ten opzichte van het Na adduct. Met de Orbitrap Fusion analyse is deze stof met directe injectie niet gedetecteerd in monster Pb. Bremerberg. In het geconcentreerde SPE extract is de stof wel goed gedetecteerd. De brutoformule is met behulp van de Orbitrap Fusion data vastgesteld op $\text{C}_{14}\text{H}_{22}\text{O}_{13}$ i.p.v. $\text{C}_{19}\text{H}_{18}\text{N}_4\text{O}_4\text{S}$ met de QToF data. In het MS^1 spectrum is duidelijk op te maken dat er geen zwavel aanwezig is maar wel een substantieel signaal van de O^{18} isotoop (zie Figuur 4-2) met een relatieve intensiteit van 2,5%.

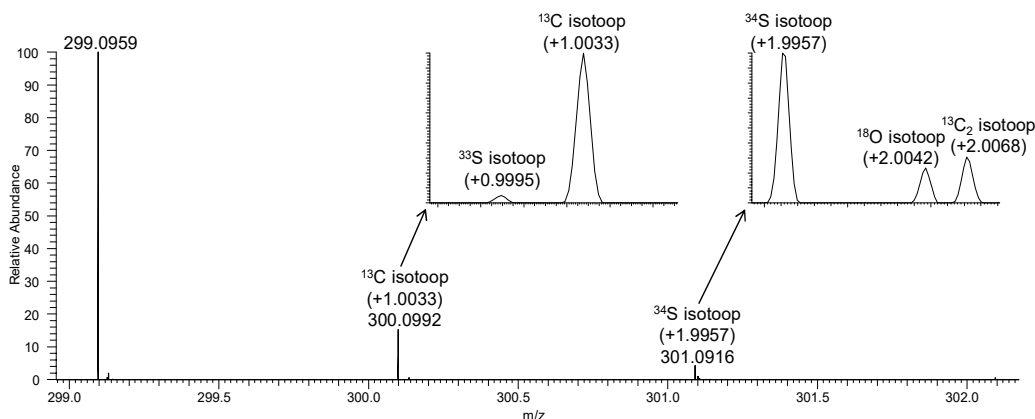


FIGUUR 4-2 FULLSCAN M+2 SPECTRUM VAN FEATURE 2431 (P10) VAN HET NH_4 ADDUCT (M/Z 416,13862) OPGENOMEN MET EEN RESOLUTIE VAN 500.000 FWHM, WAARBIJ GEEN ^{34}S ISOTOOP (MASSAVERSCHIL 1,9958) WORDT GEDETECTEERD MAAR WEL EEN RELATIEF INTENS SIGNAAL VAN ^{18}O ISOTOOP.

Deze relatief hoge intensiteit van de ^{18}O isotoop bevestigt de aanwezigheid van veel zuurstof atomen >10 in de verbinding. Vervolgens is getracht aan de hand van de brutoformule en MS^2 spectrum (zie bijlage V) de verbinding te identificeren doormiddel van *in silico* fragmentatie tools (MetFrag, Sirius en MS-Finder). Uiteindelijk heeft dit niet geleid tot goede kandidaten waardoor de identiteit van deze verbinding onbekend blijft.

Feature 3239 (N7)

Met de Orbitrap Fusion analyse is de stof met feature nr. 3239 gedetecteerd in monster Pb. Laren (reinwater) doormiddel van de directe injectie analyse evenals door de analyse van het SPE extract. Echter was de intensiteit van deze stof met m/z 299,0959 iets hoger bij de directe injectie analyse dan bij de analyse van het SPE extract, wat inhoudt dat de recovery van deze stof met de toegepaste SPE methode zeer laag is. De verdere identificatie van deze stof is daarom uitgevoerd met de data van de directe injectie analyse. Aan de hand van het MS^1 spectrum (zie Figuur 4-3) en het MS^2 spectrum (zie bijlage V) is de brutoformule vastgesteld op $C_{14}H_{20}O_5S$.

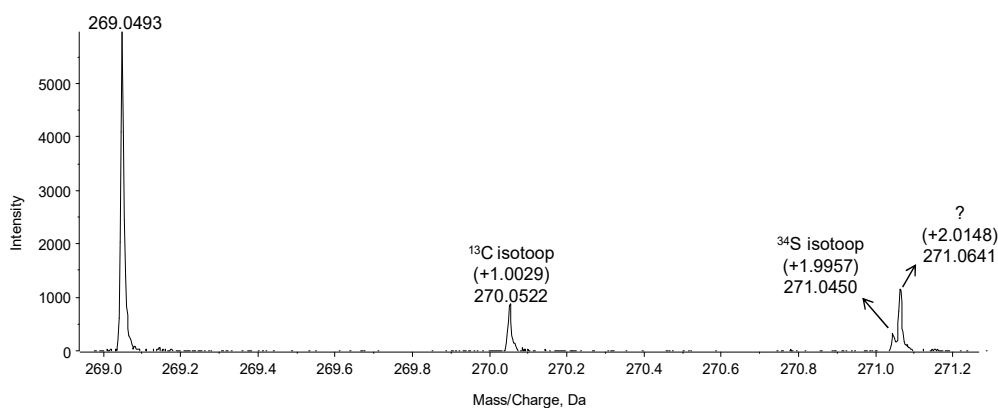


FIGUUR 4-3 FULLSCAN SPECTRUM VAN FEATURE 3239 (M/Z 299,0959) OPGENOMEN MET EEN RESOLUTIE VAN 500.000 FWHM, WAARBIJ DE ZWAVELISOTOPEN (^{33}S EN ^{34}S) EN ZUURSTOFISOTOOP (^{18}O) DUIDELIJK ZICHTBAAR ZIJN.

Vervolgens is met behulp van de brutoformule, MS^2 spectrum en *in silico* fragmentatie getracht de identiteit van de verbinding te achterhalen. De *in silico* fragmentatie leidde niet tot geschikte kandidaten (lage scores, dus weinig kans op succes). Omdat er in het chromatogram veel isomeren met dezelfde accurate massa zijn gedetecteerd is de kans groot dat het hier om natuurlijk organisch materiaal of transformatieproducten gaat. En doordat deze type verbindingen vaak niet in databases vermeld staan is de kans op een succesvolle identificatie klein.

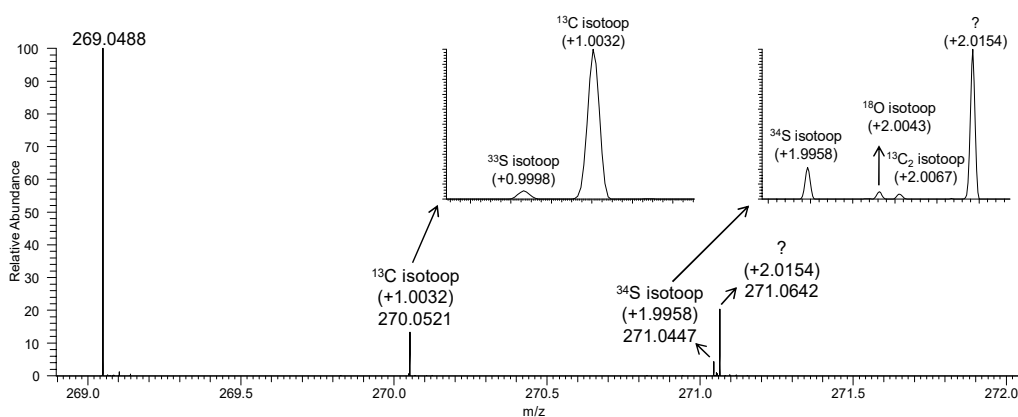
Feature 3266 (N9)

Voor de stof met feature 3266 was tijdens de identificatie van de geprioriteerde stoffen al mogelijke kandidaten gevonden. Echter was er toch twijfel eraan of de juiste brutoformule ($C_{12}H_{14}O_5S$) wel was vastgesteld, aangezien er in de QToF data een stoormassa (zie Figuur 4-4) aanwezig was op de ^{34}S isotoop positie.



FIGUUR 4-4 FULLSCAN SPECTRUM VAN FEATURE 3239 (M/Z 269,0493) OPGENOMEN MET EEN QTOF 5600+ MET EEN RESOLUTIE VAN CIRCA 30.000 FWHM, WAARBIJ ER BIJ DE VERMOEDELIJKE ZWAVELISOTOOP (^{34}S) EEN STOORMASSA (M/Z 271,0641) AANWEZIG IS.

Om er toch zeker van te zijn dat de juiste brutoformule is vastgesteld, is de stof met feature nr. 3266 geanalyseerd met de Orbitrap Fusion. Met de Orbitrap Fusion is stof gedetecteerd doormiddel van directe injectie en analyse van het SPE extract in monster Pb. Laren (ruw filter 2). Echter was de intensiteit van deze stof maar iets lager bij de directe injectie analyse dan bij de analyse van het SPE extract. De verdere identificatie van deze stof is daarom uitgevoerd met de data van de directe injectie analyse. Aan de hand van het MS^1 spectrum (zie Figuur 4-5) en het MS^2 spectrum (zie bijlage V) is de brutoformule wederom vastgesteld op $\text{C}_{12}\text{H}_{14}\text{O}_5\text{S}$.



FIGUUR 4-5 FULLSCAN SPECTRUM VAN FEATURE 3266 (M/Z 269,0488) OPGENOMEN MET EEN RESOLUTIE VAN 500.000 FWHM, WAARBIJ DE ^{34}S ISOTOOP EENVOUDIG WORDT GESCHIEDEN VAN STOORMASSA M/Z 271,0642.

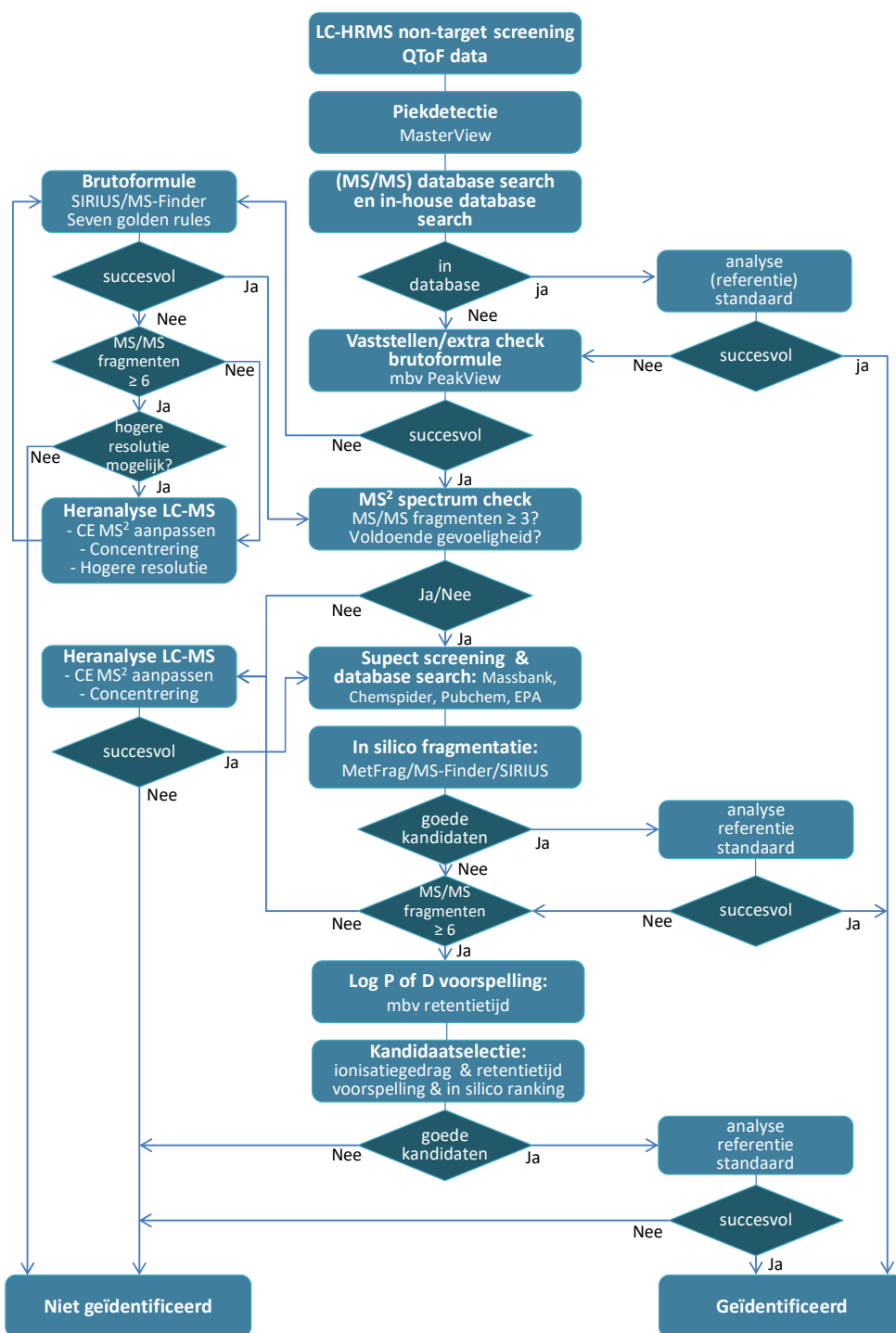
Vervolgens is met behulp van de brutoformule, MS^2 spectrum en *in silico* fragmentatie getracht de identiteit van de verbinding te achterhalen. De *in silico* fragmentatie leidde niet tot geschikte kandidaten (lage scores, met weinig kans op succes). Deze onbekende stof vertoont overeenkomsten met feature 3239, doordat beide brutoformules O_5S bevatten en zeer waarschijnlijk een sulfonzuur groep bezitten (m/z 79,9574 in MS^2 spectrum). Het is hierdoor aannemelijk dat het in dit geval ook om natuurlijk organisch materiaal of transformatieproduct gaat.

4.6.1 Conclusie Orbitrap Fusion analyse

De analyse van de vier geprioriteerde stoffen met de Orbitrap Fusion heeft niet geleid tot de complete identificatie van een van deze stoffen. De stof met feature nr. 1303 is in het geheel niet gedetecteerd. Voor de drie andere stoffen (feature nr 2431, 3239, 3266) zijn de brutoformules succesvol en met grote zekerheid vastgesteld en zijn er kwalitatief goede MS² spectra verkregen. De Orbitrap Fusion heeft duidelijk zijn meerwaarde laten zien bij het vaststellen van de brutoformule, vanwege de hogere resolutie (500.000 FWHM) die mogelijk is ten opzichte van de QToF (circa 30.000 FWHM), waardoor er meer isotopen zichtbaar zijn en er minder verstoringen optreden door achtergrondsignalen (stoormassa's).

4.7 Herziende workflow voor identificatie van onbekende verbindingen

Uit het identificatieproces van de geprioriteerde stoffen is gebleken dat een substantieel deel van de stoffen niet goed geïdentificeerd kan worden. Voor features met een goed MS² spectrum (d.w.z. > 6 fragmentionen met signaal ruis verhouding > 10) is een mogelijke oorzaak dat het om verbindingen gaat die niet opgenomen zijn in stofdatabases en spectrum bibliotheken, zoals transformatieproducten, natuurlijke stoffen of nog reeds onbekende antropogene stoffen. Daardoor kunnen deze niet gebaseerd op de MS¹ accurate massa en matches met referentiespectra nog door *in silico* fragmentatie geïdentificeerd worden [15, 22]. Maar voor veel geprioriteerde stoffen geldt dat er nog voldoende aanknopingspunten zijn voor verdere identificatie die kunnen worden verbeterd door de analysegang aan te passen. Denk hierbij aan concentreren van het monster, analyse bij een andere collision energie en analyse bij een nog hogere resolutie zoals beschreven in paragraaf 4.3. Deze mogelijkheden voor verdere identificatie zijn vaak goed uitvoerbaar en zijn daarom ook opgenomen in de doorontwikkelde workflow voor de identificatie van geprioriteerde stoffen. Zie Figuur 4-6 voor de nieuwe workflow. Het wordt dan ook aanbevolen om deze workflow toe te passen in toekomstig onderzoek voor de identificatie van onbekende organische verbindingen.



FIGUUR 4-6 WORKFLOW VOOR DE IDENTIFICATIE VOOR ONBEKENDE VERBINDINGEN

4.8 Toxicologische evaluatie van geprioriteerde stoffen

De Amerikaanse Environmental Protection Agency (U.S. EPA) heeft een database met *in vitro* toxiciteitsdata voor grote hoeveelheden stoffen en stelt die voor het publiek beschikbaar. Dit maakt high-throughput prioritering van suspects gebaseerd op hun *in vitro* toxiciteit mogelijk, mits deze in de database aanwezig zijn. Voor de Top10 mogelijke kandidaten vanuit MetFrag en MS Finder zijn CAS nummers opgezocht met hulp van het Chemical Translation Service webtool² via de ChemSpider ID (Output van MetFrag) of de InchiKey (Output van MS Finder). Vervolgens is in de database U.S. EPA gezocht naar de toxicologische gegevens van de stoffen achter deze CAS nummers. Op deze wijze kunnen toxicologische gegevens van suspects bij de features gezocht worden. Dit biedt de mogelijkheid om features toxicologisch te evalueren zoals is gedaan in een gerelateerde SPO studie [20].

TABEL 4-8. BESCHIKBAARHEID VAN TOXCAST IN VITRO TOXICITEITS DATA VOOR TENTATIVE CANDIDATES

	kandidaten per feature	kandidaten zonder CAS	kandidaten met CAS	aantal CAS	in Toxcast
p1	20	11	9	11	1
p3	15	7	8	18	1
p4	20	12	8	14	2
p9	20	5	15	24	0
n1	15	11	4	5	1
n3	15	9	6	10	0
n4	20	11	9	10	0
n5	13	12	1	1	0
n6	20	13	7	8	2
n9	20	10	10	14	2
	178	101	77	115	9

Tabel 4-8 geeft een overzicht over de aantal kandidaten per feature en de daarvoor gevonden CAS nummers in Toxcast. Omdat kandidaten meerdere keer kunnen voorkomen in de output van MetFrag en MS Finder zijn dubbelingen mogelijk. Voor minder dan de helft van de kandidaten kon een CAS nummer worden gevonden en voor slechts vijf procent van de kandidaten was in Toxcast data beschikbaar. Aan de hand van de 7 unieke CAS nummers zijn vervolgens AC50 waarden (concentratie waarbij 50% van maximale cel proliferatie is gemeten) opgezocht en minimale AC50 waarden bepaald. In Tabel 4-9 zijn deze verzameld, en de aantal van assays op die een stof getest is aangetoond. 5 van de kandidaten tonen actieve assay endpoints, relevant voor de prioritering. Deze kunnen worden besteld en met LC-HRMS gemeten om de kandidaten te bevestigen. Twee stoffen zijn kandidaten voor feature n9, namelijk Genistein en Apigenin. Deze tonen beide een lage min AC50 aan en zijn dus toxicologisch belangrijk, maar verschillen in AC50. Bij de keuze welke stoffen geprioriteerd worden voor identificatie, kan deze informatie helpen. Niettemin werd het succes van de hierboven beschreven aanpak beperkt door gebrek aan beschikbare toxicologische data voor de verschillende kandidaten van de aangetroffen features in Toxcast.

² <http://cts.fiehnlab.ucdavis.edu/conversion/batch>

TABEL 4-9. OVERZICHT CAS NUMMERS EN TOXCAST DATA PER FEATURE.

Feature	CAS number	Compound name	Min AC50	assay with min AC50	aantal assays	actieve assays	% actief
p1	672-87-7	Metyrosine	16.36815	TOX21_p53_BLA_p5_ch2	109	1	0.9
p4	3775-90-4	2-(tert-Butylamino)ethyl 2-methylprop-2-enoate	1.00E+06	inactive	116	0	0.0
n1	26225-79-6	Ethofumesate	0.194709	NVS_IC_rKATPCh	1085	52	4.8
n6	56507-37-0	Diketometribuzin	1.00E+06	inactive	240	0	0.0
n9	446-72-0	Genistein	0.012682	NVS_NR_hER	905	143	15.8
n9	520-36-5	Apigenin	0.000848	TOX21_AR_BLA_Agonist_ratio	337	82	24.3

5 Conclusies en aanbevelingen

5.1 Conclusies

Bij non-target chemische screening van kwetsbare grondwaterwinningen wordt een scala aan onbekende stoffen aangetroffen. Het is niet mogelijk alle stoffen te identificeren en toxicologisch te beoordelen. In deze studie is een prioriteringsmethodiek ontwikkeld om onbekende stoffen te prioriteren voor verdere identificatie / onderzoek op basis van voorkomen, (mogelijke) antropogene oorsprong en gedrag in de zuivering. Vervolgens is een workflow ontworpen om de meest relevante stoffen verder te onderzoeken en identificeren.

In 394 monsters van 29 kwetsbare winningen worden 612 bekende en 11087 onbekende stoffen (features) aangetroffen. Deze stoffen zijn geprioriteerd op basis van: 1) intensiteit van de respons in relatie tot een interne standaard, 2) het aantal keer dat de stof in de monsters wordt aangetroffen, 3) of de stof voorkomt in rein water en toeneemt of wordt gevormd tijdens de zuivering (mogelijk transformatieproduct gevormd tijdens zuivering of toevoeging door zuiveringsproces), 4) of de oorsprong van de stof (mogelijk) antropogeen is op basis van correlatie met winningskarakteristieken, aanwezigheid van halogenen en correlatie met bekende antropogene stoffen. Er is ook getracht een relatie te leggen met landgebruik. Omdat een sterke correlatie met landgebruik ontbrak voor de geselecteerde winningen, is dit element niet meegewogen in de prioritering. Tevens ontbrak een correlatie tussen de brutoformule en toxicologische relevantie, waardoor dit criterium eveneens niet is verwerkt in de prioritering.

Op basis van de bovenstaande prioritering zijn 12 ongeïdentificeerde stoffen in de positieve modus en 10 ongeïdentificeerde stoffen in de negatieve ionisatiemodus geselecteerd voor verder onderzoek. De op voorhand geïdentificeerde verbindingen door Vitens data zijn hier buiten beschouwing gelaten. Voor de onbekende stoffen is een workflow ontwikkeld om de meest relevante activiteiten te selecteren om deze stoffen te kunnen identificeren en (toxicologisch) evalueren. Bij toepassing van de workflow bleek dat voor 10 stoffen (45%) een beperkt aantal kandidaat stoffen konden worden aangewezen, waarvoor uiteindelijk voor 9 stoffen referentiestandaarden zijn aangeschaft. De identificatie van deze stoffen was niet succesvol, voor geen van de geprioriteerde stoffen kon door middel van referentiestandaarden de identiteit bevestigd worden. Mogelijke oorzaken voor deze lage succes rate zijn: 1) relatief lage intensiteit voor veel geprioriteerde verbindingen, waardoor er relatief veel verstoringen zijn in de verkregen MS¹ en MS² spectra en 2) dat er transformatie producten geprioriteerd zijn, die tijdens de zuivering of in de bodem ontstaan. Deze zijn vaak (nog) niet bekend en dus niet in databases te vinden. Daarnaast 3) waren van sommige kandidaat stoffen geen referentiestandaarden beschikbaar zijn en hierdoor dus ook niet in identiteit bevestigd kunnen worden. Voor deze ongeïdentificeerde stoffen en voor de resterende (55%) van de stoffen waarvoor geen prominente kandidaten konden worden gevonden is een overzicht opgesteld met aanknopingspunten voor verdere identificatie. Hiervan zijn voor vier stoffen opnieuw monsters geanalyseerd met een zeer hoge resolutie massaspectrometer. Dit heeft geleid tot de bevestiging van de vermoedelijke brutoformule voor twee stoffen en de wijziging van één veronderstelde brutoformule in een andere juiste brutoformule. De vierde stof werd niet aangetroffen in het opnieuw genomen watermonster.

In geen van de gevallen heeft heranalyse tot een complete identificatie van de onbekende stof geleid. Wel is hiermee de meerwaarde zeer hoge resolutie massaspectrometrie aangetoond, want dankzij de zeer hoge resolutie zijn meer isotopen zichtbaar en is het eenvoudiger om de juiste brutoformule vast te stellen.

Op basis van de verkregen resultaten tijdens dit onderzoek is de toegepaste identificatie workflow doorontwikkeld en is er meer nadruk gelegd op her-analyse van de te identificeren stoffen wanneer er niet genoeg MS² fragmentionen beschikbaar zijn, of wanneer er twijfel is over de vastgestelde brutoformule. Het wordt dan ook aanbevolen deze herziende workflow voor toekomstig onderzoek toe te passen.

Samenvattend blijkt uit deze studie dat de identificatie van onbekende stoffen in grondwater een complexe opgave is. Hierboven zijn verschillende activiteiten beschreven die de identificatie kunnen bevorderen. Het onderzoek naar de identiteit en de verschillende stappen blijken arbeidsintensief en lang niet altijd succesvol.

5.2 Aanbevelingen

Het project laat zien dat ook onbekende stoffen te prioriteren zijn met de beperkte beschikbare gegevens. Dit geeft richting aan vervolgonderzoek naar voorkomen en identificatie en toxicologische evaluatie van stoffen. De ontwikkelde workflow biedt handvatten om onbekende stoffen gestructureerd te beoordelen/identificeren. Het prioriteren en in het bijzonder het identificeren van een onbekende stof kost veel tijd en geld. Dit laat zien hoe belangrijk het is om op internationaal vlak samen databases uit te breiden en te delen met massa's, fragmentatiepatronen en relatieve retentietijden. Het is daarbij wel van belang dat stoffen in grondwater specifieke aandacht krijgen want het palet stoffen in grondwater verschilt van stoffen in oppervlaktewater en het meeste onderzoek richt zich op oppervlaktewater [23]. Verder is het belangrijk dat ook transformatieproducten in databases terecht komen. Bijvoorbeeld de voor water relevante database STOFF-IDENT is begonnen met het toevoegen van transformatieproducten. Tevens kunnen suspect lijsten worden aangevuld met *in silico* voorspelde transformatieproducten. Anderzijds kunnen transformatieproducten door de structurele gelijkheid met hun moederstoffen en de gelijkheid tussen MS² fragmentatiespectra geïdentificeerd worden, zoals recent beschreven [15]. De gecombineerde aanpak met zogenoemde "metabolic logic" waar naar moederstof transformatieproduct-paren met specifieke massaverschillen van transformatieprocessen worden gezocht leverde veel "onwaarschijnlijke" hits op. De hoeveelheid aan non-target screening features (11087) leidt tot een hoge kans dat toevallig features met specifieke massaverschillen worden gevonden. Extra filterstappen zouden in de toekomst transformatieproducten van toevallige hits kunnen scheiden, zoals bijvoorbeeld een beperking op transformatieprocessen die daadwerkelijk bij de toegepaste zuiveringstechnieken en omstandigheden plaatsvinden. Transformatieproducten die al in het milieu ontstaan zijn nog moeilijker te identificeren omdat Transformatieprocessen in het milieu niet altijd bekend zijn en hun producten beperkt in referentiedatabases zijn opgenomen. Met de voorgestelde aanpak krijgen deze features overigens geen hogere prioriteit tenzij hun intensiteit tijdens de zuivering (verder) toe neemt.

Complete identificatie van alle (relevante) onbekende stoffen is op dit moment nog niet haalbaar. Het is daarom noodzakelijk om alternatieve benaderingen te verkennen om de waterkwaliteit te beoordelen zonder dat de identiteit van (alle) stoffen bekend is. Bioassays kunnen bijvoorbeeld helpen om een scala aan specifieke effecten te meten. Deze assays kunnen een oordeel vellen over de waterkwaliteit voor verschillende toxicologische eindpunten. Daarnaast kunnen de effecten ook statistisch worden gerelateerd aan de aangetroffen (onbekende) stoffen, waardoor complete identificatie niet altijd nodig is voor het opzetten van een monitoringprogramma en een robuuste risicobeoordeling.

Tevens kunnen zogenoemde effect gestuurde analyse (Effect Directed Analysis) methoden worden ingezet waarbij monster worden gefractioneerd met behulp van chromatografische scheiding en deze fracties individueel toxicologisch en chemisch worden geanalyseerd om te bepalen welke stoffen verantwoordelijk zijn voor de waargenomen effecten [24]. Op deze wijze kunnen onbekende stoffen die (mogelijk) een bepaald toxicologisch effect te weeg brengen worden geselecteerd voor aanvullende monitoring of identificatie en kan sneller een oordeel over de waterkwaliteit worden geveld. De in deze studie beschreven non-target methoden kunnen in combinatie met effect analyses worden ingezet om het risico gestuurde monitoring programma van drinkwater bedrijven vorm te geven. Zo kan risico en probleemgestuurd worden gemonitord, zoals dat in de recente wijzigingen van annex II van het drinkwaterbesluit van de drinkwaterbedrijven wordt verlangd.

6 Literatuur

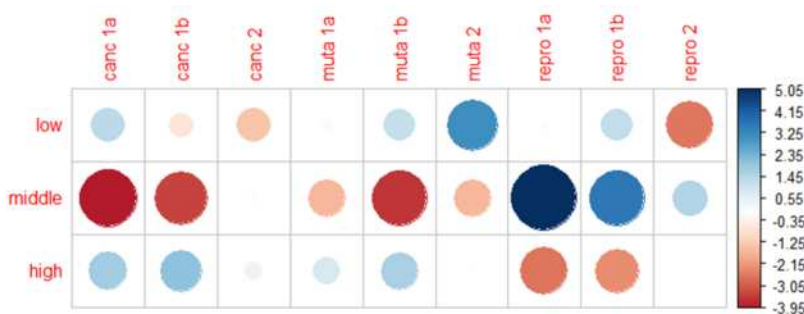
1. Zedda, M. and C. Zwiener, *Is nontarget screening of emerging contaminants by LC-HRMS successful? A plea for compound libraries and computer tools*. Analytical and Bioanalytical Chemistry, 2012. **403**(9): p. 2493-2502.
2. Bletsou, A.A., et al., *Targeted and non-targeted liquid chromatography-mass spectrometric workflows for identification of transformation products of emerging pollutants in the aquatic environment*. TrAC - Trends in Analytical Chemistry, 2015. **66**: p. 32-44.
3. ter Laak, T., et al., *Literature survey of transformation products in the water cycle*. 2015, KWR: Nieuwegein. p. 65.
4. Lambropoulou, D.A. and L.M.L. Nollet, *Transformation Products of Emerging Contaminants in the Environment*. Vol. 1. 2014, Chichester, UK: Wiley. 1-383
5. Van Leerdam, J.A., et al., *Brede screening van drinkwater: op zoek naar onbekende stoffen*. H2O online, 2015. **november / december**.
6. Van Leerdam, J.A., et al., *Brede chemische screening voor het monitoren van de waterkwaliteit*. H2O online, 2015. **Oktober**.
7. van Leerdam, J.A., et al., *Implementatie van LC-MS screeningstechnieken voor monitoring waterkwaliteit*. 2015, KWR, Watercycle Reserach Institute: Nieuwegein, the Netherlands.
8. De Jonge, M., *Kwetsbaarheid Vitens winningen*. 2016, Vitens: Zwolle, The Netherlands. p. 17.
9. Stuyfzand, P. and R. Stuurman, *Elf verschillende bronnen van verzilting van grondwater in Nederland*. H2O, 1996. **29**: p. 374-377.
10. Menz, J., et al., *Transformation products in the water cycle and the unsolved problem of their proactive assessment: A combined in vitro/in silico approach*. Environment International, 2017. **98**: p. 171-180.
11. Schollée, J.E., et al., *Prioritizing Unknown Transformation Products from Biologically-Treated Wastewater Using High-Resolution Mass Spectrometry, Multivariate Statistics, and Metabolic Logic*. Analytical Chemistry, 2015. **87**(24): p. 12121-12129.
12. Cui, X. and G.A. Churchill, *Statistical tests for differential expression in cDNA microarray experiments*. Genome Biology, 2003. **4**(4).
13. Sleno, L., *The use of mass defect in modern mass spectrometry*. Journal of Mass Spectrometry, 2012. **47**(2): p. 226-236.
14. Nagao, T., et al., *Power of isotopic fine structure for unambiguous determination of metabolite elemental compositions: In silico evaluation and metabolomic application*. Analytica Chimica Acta, 2014. **813**: p. 70-76.
15. Schollée, J.E., et al., *Similarity of High-Resolution Tandem Mass Spectrometry Spectra of Structurally Related Micropollutants and Transformation Products*. Journal of the American Society for Mass Spectrometry, 2017. **28**(12): p. 2692-2704.
16. Reemtsma, T., et al., *Mind the gap: persistent and mobile organic compounds - water contaminants that slip through*. Environmental Science & Technology, 2016. **asap**: p. asap.
17. Sjerps, R.M.A., et al., *Data-driven prioritization of chemicals in various water types using suspect screening LC-HRMS*. Water Research, 2016: p. 254-264.
18. van Leerdam, J.A., et al., *Exploring the boundaries of non-target screening with Liquid Chromatography coupled to ESI-MS*. 2017, KWR Watercycle Reserach Institute: Nieuwegein, The Netherlands. p. 89.
19. Mahieu, N.G., et al., *Defining and Detecting Complex Peak Relationships in Mass Spectral Data: The Mz.unity Algorithm*. Analytical Chemistry, 2016. **88**(18): p. 9037-9046.

20. Rosa M.A. Sjerps, et al., *Clustering and prioritisation to design a risk based monitoring program in groundwater sources for drinking water*. 2018, KWR Watercycle Research Institute: Nieuwegein, The Netherlands. p. 35.
21. Leerdam, J., et al., *Bevestiging van de identiteit van geprioriteerde suspects*. 2017, KWR Watercycle Research Institute: Nieuwegein, The Netherlands. p. 37.
22. Bletsou, A.A., et al., *Targeted and non-targeted liquid chromatography-mass spectrometric workflows for identification of transformation products of emerging pollutants in the aquatic environment*. *TrAC Trends in Analytical Chemistry*, 2015. **66**: p. 32-44.
23. ter Laak, T.L., et al., *Broad target chemical screening approach used as tool for rapid assessment of groundwater quality*. *Science of the Total Environment*, 2012. **427-428**: p. 308-313.
24. Brack, W., *Effect-directed analysis: A promising tool for the identification of organic toxicants in complex mixtures?* *Analytical and Bioanalytical Chemistry*, 2003. **377(3)**: p. 397-407.

7 Bijlagen

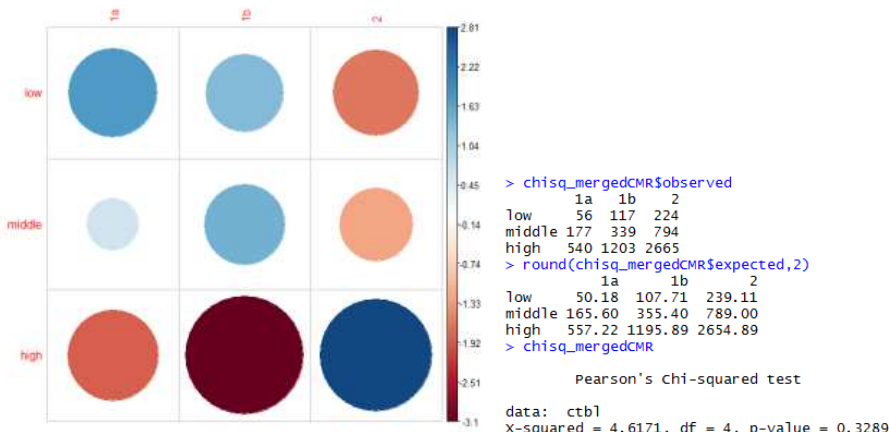
Bijlage I Correlaties tussen CMR classificaties en op bruto formule gebaseerde toxiciteitsklassen van CMR geclassificeerde stoffen uit ECHA.

Correlaties tussen uitgesplitste (boven) en overkoepelende (onder) CMR classificaties en op bruto formule gebaseerde toxiciteitsklassen van CMR geclassificeerde stoffen (ECHA). De grootte van de cirkel weerspiegelt de relatieve bijdrage. De kleur van de cirkel geeft aan of het een positieve (blauwtinten) of negatieve (roodtinten) associatie betreft, waarbij de grootte van deze associatie overeenkomt met de intensiteit van de kleur.



```
> chisq$observed
      canc 1a  canc 1b  canc 2  muta 1a  muta 1b  muta 2  repro 1a  repro 1b  repro 2
low         27         35         69          2         20         83          27         62         72
middle      36         91        271          2         25        172         139        223        351
high       258        479        918         24        192        656         258         532       1091
> round(chisq$expected,2)
      canc 1a  canc 1b  canc 2  muta 1a  muta 1b  muta 2  repro 1a  repro 1b  repro 2
low        20.84   39.28   81.67    1.82   15.39   59.14    27.53   53.04   98.29
middle     68.77  129.61  269.50    6.00   50.77  195.16   90.83  175.02  324.34
high     231.39  436.11  906.83   20.18  170.84  656.69  305.64  588.93 1091.37
> chisq
```

Pearson's Chi-squared test
 data: tbl
 X-squared = 134.02, df = 16, p-value < 2.2e-16



Bijlage II Overzicht referentie-standaarden voor identificatie van kandidaat stoffen

REFERENTIESTANDAARDEN KANDIDAATSTOFFEN POSITIEVE ANALYSE

P	Feature Nr	Stofnaam	Cas nr	Leverancier	Hoeveelheid	Artikel Nr	Prijs
1	1504	Metirosine N-Methyltyrosine	672-87-7 537-49-5	Sigma-Aldrich Sigma-Aldrich	100mg 50mg	M8131-100MG 04632-50MG	97€ 85€
3	607	2-Aminodecanedioic acid N-(tert-butoxycarbonyl)valine	3721-88-8 54895-12-4	Niet leverbaar Sigma-Aldrich	nvt 25g	nvt 15528-25G	nvt 49.8€
4	525	(1-Isopropyl-4-piperidinyl)acetic acid	754183-67-0	SCBT*	100mg	sc-282132	61€
9	2554	3-Anilino-1-propanesulfonic acid 3-(Ethylamino)-4methylbenzenesulfonic acid	72943-20-5 98-40-8	Sigma-Aldrich Sigma-Aldrich	250mg 1g	S562688-250MG S576875-1G	125€ 125€

* SCBT = Santa Cruz Biotechnology

REFERENTIESTANDAARDEN KANDIDAATSTOFFEN NEGATIEVE ANALYSE

N	Feature Nr	Stofnaam	Cas nr	Leverancier	Hoeveelheid	Artikel Nr	Prijs
1	1401	Ethofumesate	26225-79-6	Sigma-Aldrich	250mg	45479-250MG	68.5€
3	1350	Ferulic acid 4-sulfate	86321-29-1	TRC Canada	100mg	F308920	150\$
4	1442	Methyl 2-methoxy-5-(methylsulfonyl)benzoate 3-[(4-Methoxyphenyl)sulfonyl]propanoic acid 2-[4-(Methoxycarbonyl)phenyl]ethanesulfonic acid	63484-12-8 91062-23-6 756469-30-4	TCI SCBT* Niet leverbaar	5g 500mg nvt	M1963 sc-309887	61€ 291€
5	1972	3-[(3,4-Dimethoxyphenyl)sulfonyl]propanoic acid	100059-84-5	VitasM	5mg	STK803172	-
6	475	metribuzin-diketo	56507-37-0	LGC	5mg	DRE-C15200700	104€
9	3266	[(2S)-5-Oxotetrahydro-2-furanyl]methyl 4-methylbenzenesulfonate Genistein Apigenin	58879-34-8 446-72-0 520-36-5	Sigma-Aldrich Sigma-Aldrich Sigma-Aldrich	1g 10 mg 25 mg	364568-1G 92136-10MG 10798-25MG	145€ 98.5€ 92€

* SCBT = Santa Cruz Biotechnology

Bijlage III Mogelijke transformatie producten geprioriteerde stoffen

OVERZICHT MOGELIJKE TRANSFORMATIEPRODUCTEN MET MOEDERSTOFFEN (POSITIEVE ANALYSE)

P	Feature Nr	Componentnaam	Mogelijk transformatieproduct	Mogelijke moederstof
1	1504	196.0965 / 11.83	Oxidatie (+2O), Oxidatie (+2O)	164.1066 / 20.73, 164.1066 / 20.73
2	746	284.0974 / 11.16	Oxidatie (+2O)	252.1069 / 11.93
3	607	218.1380 / 10.53	OA_Oxidatie (+2O)	222.1232 / 17.53, 186.1476 / 19.45
4	525	186.1477 / 12.30	Desamino (-HN)	201.1590 / 14.71
5	1286	341.1219 / 18.29	Oxidatie (+3O)	293.1365 / 21.83
6	354	469.3140 / 36.98	-	-
7	1303	557.2218 / 21.63	-	-
8	548	198.1340 / 12.97	-	-
9	2554	216.0674 / 10.03	Oxidatie (+O)	200.0733 / 22.45
10	2431	421.0951 / 12.61	Oxidatie (+4O)	357.1161 / 15.66
11	1207	445.2680 / 14.50	Oxidatie (+2O), Oxidatie (+2O)	164.1066 / 20.73, 164.1066 / 20.73
12	706	261.1221 / 17.57	Oxidatie (+2O)	252.1069 / 11.93

OVERZICHT MOGELIJKE TRANSFORMATIEPRODUCTEN MET MOEDERSTOFFEN (NEGATIEVE ANALYSE)

N	Feature Nr	Componentnaam	Mogelijk transformatieproduct	Mogelijke moederstof
1	1401	285.0802 / 21.16	acetylation + C2H2O, acetylation + C2H2O, Desethyl (-C2H2, Desethyl (-C2H2, Desethyl (-C2H2, Desethyl (-C2H2, acetylation + C2H2O, sulfonation + SO3, acetylation + C2H2O	243.0695 / 40.53, 243.0695 / 40.53, 313.1122 / 26.11, 313.1122 / 26.11, 313.1113 / 25.17, 313.1107 / 40.39, 243.0699 / 39.83, 205.1232 / 27.15, 243.0695 / 40.53
2	911	230.0616 / 10.48	Oxidatie (+O), Oxidatie (+O)	214.0631 / 25.69, 214.0631 / 25.69
3	1350	273.0089 / 13.93	Oxidatie (+4O), sulfonation + SO3, sulfonation + SO3	209.0275 / 19.77, 193.0501 / 19.69 193.0501 / 19.69
4	1442	243.0336 / 12.23	Desethyl (-C2H2, Desethyl (-C2H2	271.0641 / 23.49, 271.0646 / 24.23
5	1972	273.0450 / 17.29	Oxidatie (+O), Oxidatie (+O), Oxidatie (+3O), Oxidatie (+O), Desethyl (-C2H2, Oxidatie (+3O), Desethyl (-C2H2, Oxidatie (+O), Oxidatie (+O), Desethyl (-C2H2	257.0489 / 22.52, 257.0481 / 21.41, 225.0597 / 24.31, 257.0489 / 22.52, 301.0726 / 25.37, 225.0591 / 23.75, 301.0736 / 26.03, 257.0473 / 24.25, 257.0481 / 21.41, 301.0726 / 25.37
6	475	183.0896 / 13.33	-	-
7	3239	299.0952 / 23.21	-	-
8	1750	271.0653 / 17.51	Desethyl (-C2H2, Desethyl (-C2H2, Desethyl (-C2H2, Desethyl (-C2H2, Oxidatie (+O), Desethyl (-C2H2, sulfonation + SO3	299.0953 / 26.56, 299.0970 / 27.75, 299.0970 / 27.75, 299.0953 / 26.56, 255.0724 / 21.48, 299.0953 / 26.56, 191.1070 / 26.21
9	3266	269.0484 / 24.14	Desethyl (-C2H2	297.0801 / 26.24
10	1948	253.0552 / 21.78	Waterafsplitsing (-H2O), Waterafsplitsing (-H2O), Desethyl (-C2H2, Desethyl (-C2H2, Waterafsplitsing (-H2O), Waterafsplitsing (-H2O)	271.0636 / 22.86, 271.0641 / 23.49, 281.0857 / 24.36, 281.0857 / 24.36, 271.0646 / 24.23, 271.0646 / 24.23

Bijlage IV Analyse methode non-target screening Orbitrap Fusion

Monstervoorbewerking vaste fase extractie

Voor de monstervoorbewerking is 500 mL monster overgebracht in een erlenmeyer en is daarna aangezuurd tot pH 2,3 met behulp van zoutzuur. Vervolgens is het monster met een vaste fase kolom geïsoleerd (OASIS HLB, 200 mg, glass, 6 cc) verkregen van Waters (Etten-Leur, Nederland). Aansluitend is de vaste fase kolom voor 1 uur gedroogd met lucht waarna de elutie is uitgevoerd met 7.5 mL van 8:2 (v/v) acetonitril/methanol. Daarna is het eluaat verdampt met behulp van een Barkey optocontrol (Leopoldshöhe, Duitsland) doormiddel van een zachte stikstoftlow met een temperatuur van circa 75 °C (blok temperatuur van 300 °C) tot een volume van 250 µL was bereikt. Daarna is 750 µL ultrapuur water aan het extract toegevoegd met daarin 13.333 µg/L interne standaard oplossing (bevat atrazine-d5, bentazon-d6, fenuron en neburon). Aansluitend is het extract gefiltreerd met een 0.2 µm Phenomenex Phenex geregenereerde cellulose filter (Utrecht, Nederland) en is vervolgens overgebracht in een autosampler vial van 1,8 mL voor analyse met de Orbitrap Fusion.

Monstervoorbewerking directe injectie

Vijftig mL monster is overgebracht in een maatkolf van 50 mL hieraan is 1,0 µg/L interne standaard toegevoegd (bevat atrazine-d5, bentazon-d6, fenuron en neburon). Vervolgens is het monster gefiltreerd met een 0.2 µm Phenomenex Phenex geregenereerde cellulose filter (Utrecht, Nederland) en is daarna overgebracht in een autosampler vial van 1,8 mL voor analyse met de Orbitrap Fusion.

Orbitrap Fusion non-target screening methode

Instellingen Vanquish UHPLC , autosampler en kolomoven:

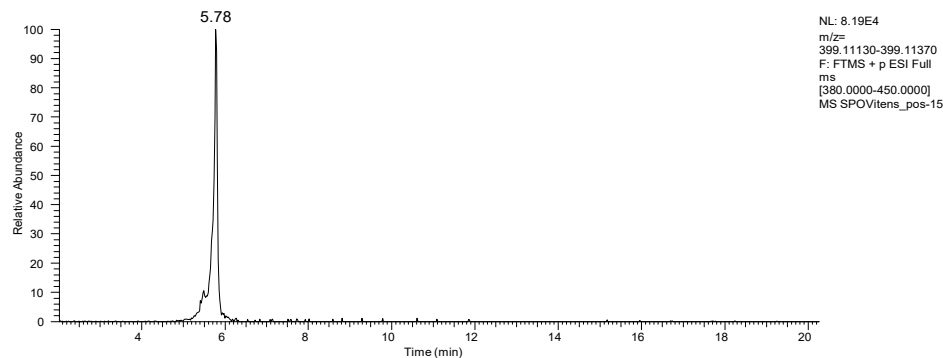
- Kolom: Waters Xbridge C18 XP, 2.1 x 150 mm, 2.5 µm
- Mobiele fase A: ultrapuur water + 0.05% mierenzuur
Mobiele fase B: acetonitril + 0.05% mierenzuur
- Gradiënt: 5% B voor 1 minuut. Dan lineair van 5% B naar 100% B in 24 min. Dan 100% B voor 4 minuten. Vervolgens terug naar begin condities in 0,5 min, en 5% B voor 4,5 min.
- Flow: 250 µL/min // Injectie volume: 100 µL (directe injectie) en 20 µL (injectie vaste fase extract)
- kolomoven temperatuur: 25°C

Instellingen Orbitrap Fusion massaspectrometer:

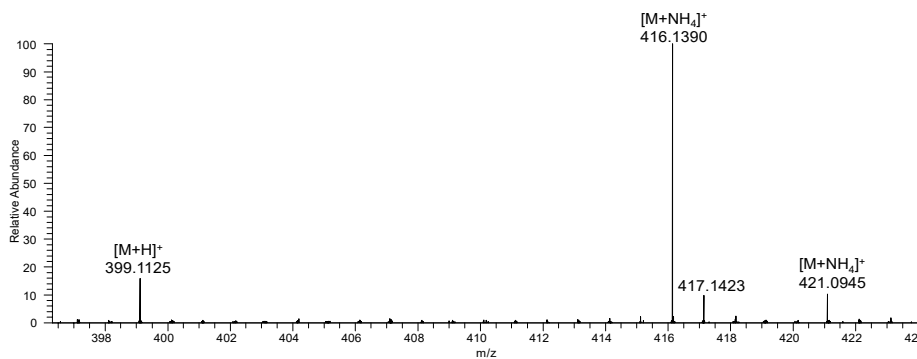
- Resolutie: 240.000 of 500.000 FWHM
- Mass accurateid < 2 ppm
- Massagebied Full scan: 100-1000 Da
- Massagebied MS² scan: 50-550 Da
- Ionisatie: positieve en negatieve mode
- Bron: electrospray (ESI)
- Vaporizer: 300°C
- Spray voltage: 3000 volt (positief) en 2500 volt (negatief)
- Sheat gas: 40 arbitraire eenheden
- Auxillary gas: 10 arbitraire eenheden
- Sweep gas: 5 arbitraire eenheden
- Acquisitie van één MS² spoor per analyse van m/z 399,11 of m/z 299,10 of m/z 269,05

Bijlage V Chromatogrammen en massaspectra Orbitrap Fusion analyse

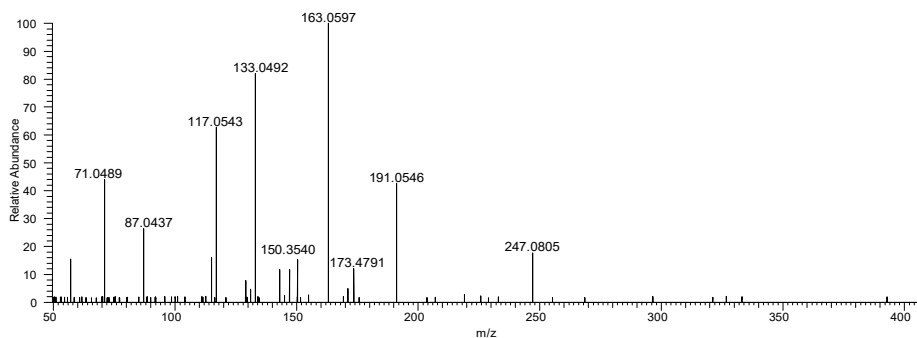
Resultaten Orbitrap Fusion analyse feature 2431 (P10)



XIC van m/z 399,1125 (feature 2431) van het geconcentreerde SPE extract van Pb. Bremerberg.

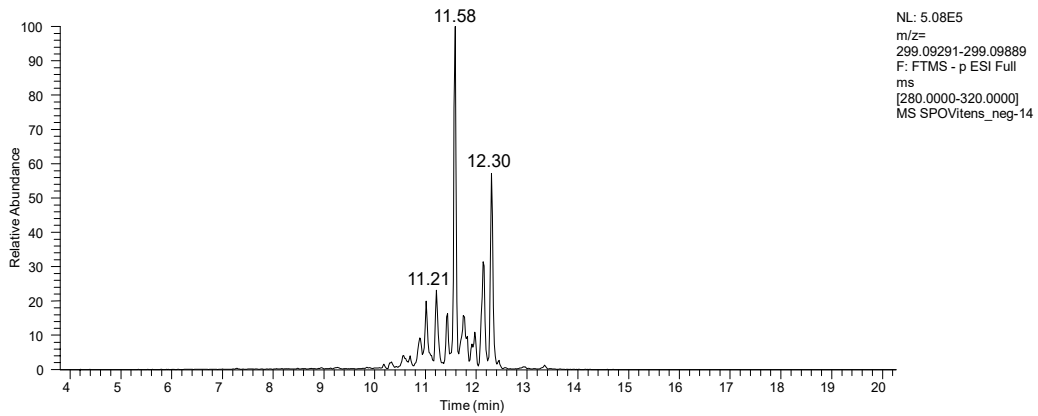


Fullscan spectrum van het geconcentreerde SPE extract van Pb. Bremerberg.

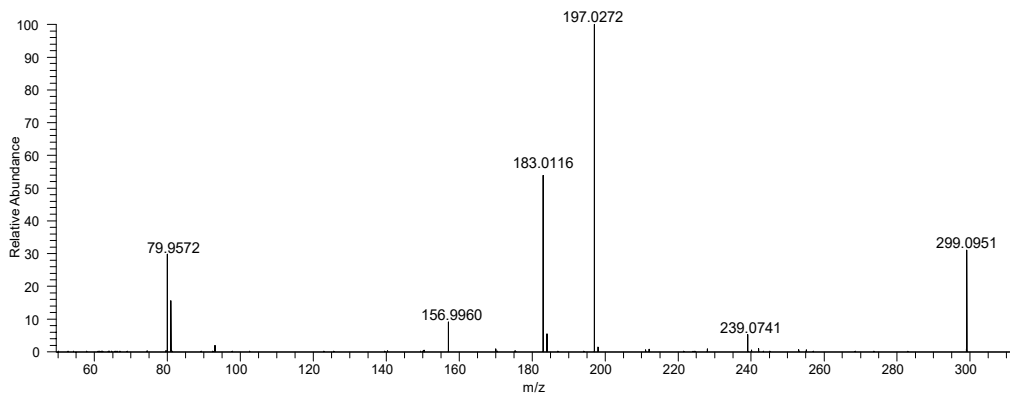


MS² spectrum (precursor 399,11 met HCD 25%) van het geconcentreerde SPE extract van Pb. Bremerberg.

Resultaten Orbitrap Fusion analyse feature 3239 (N7)

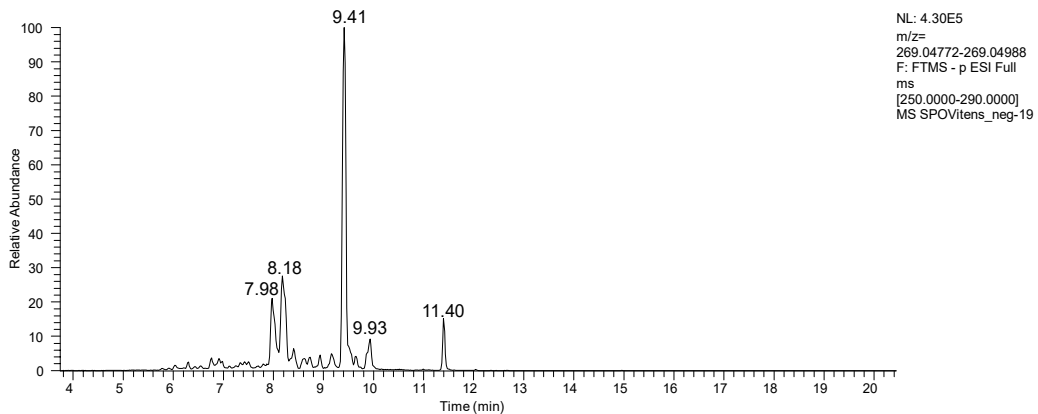


XIC van m/z 299,0959 (feature 3239) van het direct injectie monster Pb. Laren (reinwater).

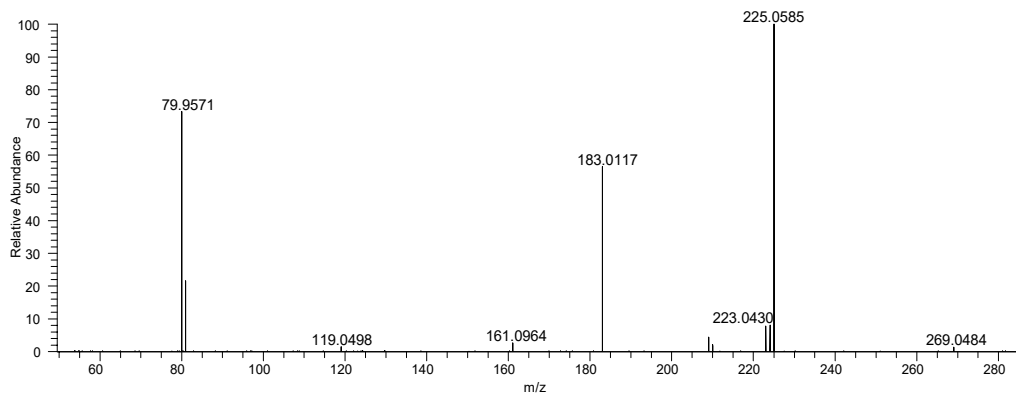


MS² spectrum (precursor 299,10, HCD 45%, RT 11.58) van het direct injectie monster Pb. Laren (reinwater).

Resultaten Orbitrap Fusion analyse feature 3266 (N9)



XIC van m/z 269,0488 (feature 3266) van het direct injectie monster Pb. Laren (ruw filter 2).



MS² spectrum (precursor 269,05, HCD 45%, RT 9.41) van het direct injectie monster Pb. Laren (ruw filter 2).