

Bedrijfstakonderzoek
BTO 2020.056 | November 2020

**Voorspelling en
validatie van de
verwijdering van
organische
microverontreinigingen
uit water; deel 1:
stofselectie**

Bedrijfstakonderzoek

KWR

Bridging Science to Practice

Rapport

Voorspelling en validatie van de verwijdering van organische microverontreinigingen uit water; deel 1: stofselectie

BTO 2020.056 | November 2020

Dit onderzoek is onderdeel van het collectieve Bedrijfstakonderzoek van KWR, de waterbedrijven en Vewin.

Opdrachtnummer

402045/030

Projectmanager

Erwin Beerendonk

Opdrachtgever

BTO - Bedrijfsonderzoek

Auteurs

Roberta Hofman-Caris, Bas Wols, Dirk Vries, Martin Korevaar, Wolter Siegers

Kwaliteitsborger

Emile Cornelissen

Verzonden naar

Dit rapport is verspreid onder BTO-participanten.

Een jaar na publicatie is het openbaar.

Keywords

microveorntreinigingen, QSPRs, drinkwaterzuivering, actieve kool, (geavanceerde) oxidatie, membraanfiltratie

Jaar van publicatie

2020

Meer informatie

dr. ir. C.H.M. Hofman-Caris

T 0653198167

E roberta.hofman-caris@kwrwater.nl

PO Box 1072

3430 BB Nieuwegein

The Netherlands

T +31 (0)30 60 69 511

F +31 (0)30 60 61 165

E info@kwrwater.nl

I www.kwrwater.nl

KWR

November 2020 ©

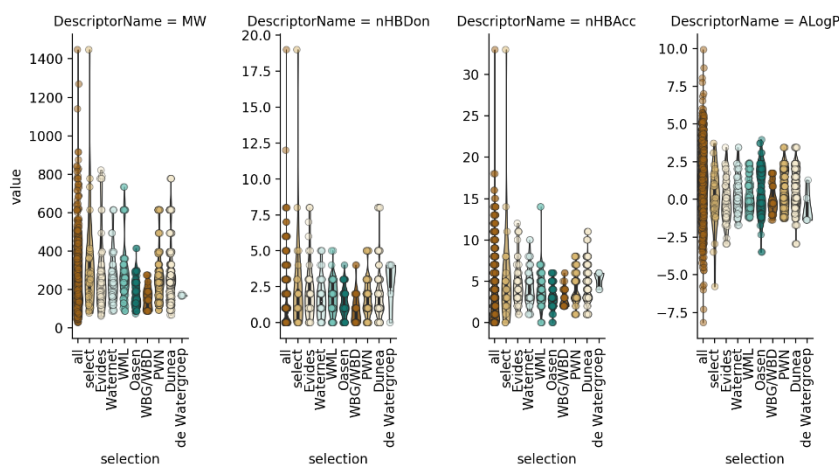
Alle rechten voorbehouden aan KWR. Niets uit deze uitgave mag - zonder voorafgaande schriftelijke toestemming van KWR - worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen,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

Managementsamenvatting

Wiskundige selectiemethode resulteert in lijst modelstoffen voor onderzoek naar (drink)waterzuiveringstechnieken

Auteurs dr.ir. Roberta Hofman-Caris, dr.ir. Bas Wols, dr.ir. Dirk Vries, dr.ir. Martin Korevaar, Wolter Siegers.

Met een objectieve, wiskundige selectiemethode is een lijst modelstoffen vastgesteld die verontreinigingen in de (drink)waterzuivering zo goed mogelijk vertegenwoordigen. Voor de set van 30 componenten die hieruit naar voren kwam is een methode ontwikkeld waarmee ze zoveel mogelijk allemaal tegelijkertijd zijn te analyseren. Op deze manier kunnen modellen voor adsorptie-, membraan- en oxidatieprocessen in de drinkwaterzuivering worden gevalideerd en vergeleken. Het valideren van modellen met praktijkgegevens is belangrijk omdat dit de betrouwbaarheid ervan vergroot en de noodzaak tot kostbaar experimenteel onderzoek terugdringt. Indien door middel van modellering van zuiveringsprocessen kan worden voorspeld hoe een bepaalde verontreiniging zich zal gedragen, kan veel gericht onderzoek worden gedaan, of zal experimenteel onderzoek in sommige gevallen zelfs niet meer nodig zijn.



“Violoplots” van een aantal eigenschappen die in verschillende sets componenten voorkomen

Belang: representatieve set microverontreinigingen voor modelonderzoek

Modellen die gebaseerd zijn op QSPRs (quantitative structure property relationships) voorspellen hoe een bepaalde organische microverontreiniging zich in een (drink)waterzuivering zal gedragen. Dit maakt onderzoek goedkoper, omdat gericht kan worden gezocht naar de meest veelbelovende technieken, of omdat zelfs zonder experimenten reële voorspellingen mogelijk zijn. Het is wel van cruciaal belang dat de gebruikte modellen betrouwbaar en

thus gevalideerd zijn. Hiervoor zijn experimenten nodig die in kaart brengen welke moleculaire eigenschappen van stoffen van invloed zijn op verwijderings- en zuiveringsprocessen in drinkwater. Dit is mogelijk met modelstoffen, die deze eigenschappen zo goed mogelijk bestrijken. In dit onderzoek is een methode ontwikkeld om een set modelstoffen te kunnen samenstellen die met zo min mogelijk analysemethoden kan worden geanalyseerd. Dit werkt kostenbesparend

Aanpak: Objectieve wiskundige selectiemethode

Vanwege de jarenlange ervaring in zuiveringstechnieken en modellering beschikt KWR over moleculaire parameters en 'vingerafdrukken' van een omvangrijke reeks van 2.400 stoffen. Aan de hand van expert judgement is op grond van deze kennis een selectie gemaakt van de parameters die waarschijnlijk een rol spelen bij adsorptie-, membraan- en oxidatieprocessen in de drinkwaterzuivering. Daarna is gekeken welke range deze parameters bestrijken bij veel voorkomende verontreinigingen in bronnen voor drinkwater. Met behulp van objectieve wiskundige technieken is vervolgens een selectie gemaakt van 30 componenten, die representatief zijn voor al deze verontreinigingen. Met deze informatie als basis is een analysemethode ontwikkeld voor het gelijktijdig analyseren van de meeste stoffen. Daardoor is slechts een beperkt aantal analysemethoden nodig. Als uitkomst kan de validatie van modellen die het effect van waterzuiveringstechnieken voorspellen, betrouwbaar en goedkoop worden uitgevoerd en zijn de verkregen resultaten representatief voor de praktijksituatie.

Resultaten: Lijst van representatieve microverontreinigingen

Uit dit onderzoek komt een lijst met modelstoffen voort, representatief voor de hele range aan moleculaire eigenschappen zoals aangetroffen bij verontreinigingen in drinkwaterbronnen die de effectiviteit van verschillende zuiveringstechnieken beïnvloeden. Met deze lijst kunnen voorspellende modellen worden gevalideerd die aangeven in welke mate bepaalde componenten uit (drink)water zijn te verwijderen met verschillende technieken

Toepassing: Validatie van zuiveringsmodellen

De samengestelde lijst van modelstoffen wordt gebruikt voor de validatie van QSPR-modellen. Deze modellen voorspellen hoe goed een bepaalde component met bepaalde techniek te verwijderen is. Met de resultaten van dit project kunnen betrouwbare modellen worden ontwikkeld die inzicht geven in de zuivering van bekende maar ook van nieuwe verontreinigingen.

Rapport

Dit onderzoek is beschreven in het rapport *Voorspelling en validatie van de verwijdering van organische microverontreinigingen uit water* (BTO-2020.056)

Meer informatie

dr. ir. C.H.M. Hofman-Caris
T 0653198167
E roberta.hofman-caris@kwrwater.nl

PO Box 1072
3430 BB Nieuwegein
The Netherlands



Bedrijfsparagraaf

In de bedrijfsparagraaf wordt de relevantie van het onderzoek voor de drinkwaterbedrijven beschreven. Kennis en inzicht in de verwijdering van organische microverontreinigingen is om meerdere redenen relevant voor drinkwaterbedrijven:

Nieuwe stoffen in het nieuws

Er is een toename van OMV in bronnen voor drinkwater. Door verbeterende analysetechnieken worden deze in steeds lagere concentraties gedetecteerd. De laatste jaren is er bijna elk jaar wel een “nieuwe” microverontreiniging die het nieuws haalt of in ieder geval in een deel van waterwereld voor opschudding zorgt. Op dergelijke momenten worden er direct dure meetcampagnes gestart om te kunnen uitzoeken of, en in welke mate, een dergelijke stof door onze zuiveringen wordt verwijderd en dus of deze microverontreiniging uiteindelijk in het drinkwater terechtkomt. Het zou een geweldige stap vooruit zijn als we op basis van stoffeigenschappen al direct met een grote mate van waarschijnlijkheid kunnen zeggen in welke hoeverre een dergelijke microverontreiniging door onze zuiveringen wordt verwijderd. Dit onderzoek draagt bij aan de ontwikkeling van die tool.

Keuze over innamestops of bijschakelen zuivering

Voor onbekende nieuwe stoffen is het lastig in te schatten wanneer we dan daadwerkelijk extra zuiveringstechnieken in moeten zetten, en ook hoe goed die dan werken. We hebben behoefte aan een duidelijke, volledige en nauwkeurige voorspellingstool. Dit kan helpen bij de inschatting hoe goed de aanwezige processen werken, of een extra zuiveringstechniek moet worden toegepast, of dat het beter is over te gaan tot een innamestop. Dit is relevant voor de drinkwaterbedrijven of we de inname MOETEN staken, of KUNNEN staken.

Voorkómen, verwijderen, chemisch omzetten

Bescherming van de bron is belangrijk om ervoor te zorgen dat ongewenste stoffen zo min mogelijk aanwezig zijn in het oppervlakte- of grondwater bestemd voor de productie van drinkwater. Als dat niet of in beperkte mate mogelijk is en de stoffen toch in het ruwwater aanwezig zijn, wordt door veel bedrijven ingezet op verwijdering van de desbetreffende stoffen. Alleen wanneer ook verwijdering (technisch en/of economisch) onvoldoende mogelijk is, zal overwogen worden zuiveringsstappen in te zetten die deze stoffen omzetten. Deze aanpak (waar mogelijk stoffen niet omzetten, maar inzetten op bronbescherming en verwijdering) wordt gehanteerd uit voorzorg, omdat het bij omzetting mogelijk is ongewenste (en onbekende) stoffen te vormen met toxische eigenschappen.

Voorzorgsprincipe

Vanuit het voorzorgsprincipe is het van belang om te weten wat er op ons afkomt: wat er niet in onze bronnen komt hoeven we er ook niet uit te halen, en hoe eerder we inzicht hebben wat er op ons afkomt des te meer tijd we hebben om te kunnen reageren. Indien de zuivering uitgebreid moet worden, is het van belang om te weten hoe de verschillende zuiveringsprocessen de relevante stoffen kunnen omzetten. Een voorspellingsmodel voor verwijdering van (organische) microverontreinigingen met verschillende zuiveringstechnieken in diverse watermatrices kan de drinkwaterbedrijven hierbij ondersteunen.

Samenvatting

Bronnen voor drinkwater bevatten steeds meer (organische) microverontreinigingen, doordat er steeds meer stoffen worden geproduceerd, en uiteindelijk geloosd. Experimenteel onderzoek naar effectieve en efficiënte zuiveringsmethoden is kostbaar. Daarom is het van belang goede modellen te ontwikkelen, waarmee het effect van een bepaalde zuiveringsmethode voor een microverontreiniging kan worden voorspeld. Aan de hand van betrouwbare voorspellingen kan het onderzoek worden gericht op een specifiek, veelbelovend zuiveringsproces, of zouden experimenten met een nieuw type microverontreiniging zelfs overbodig kunnen worden. Dit betekent een forse kostenbesparing voor de drinkwaterbedrijven. “Quantitative structure activity relationships” (QSARs) of “quantitative structure property relationships” (QSPRs) kunnen worden gebruikt bij het opzetten van dergelijke modellen. Hierbij wordt een statistisch verband gezocht tussen bepaalde moleculaire eigenschappen van stoffen en het gedrag dat ze vertonen. Modellen gebaseerd op QSPRs zijn veelbelovend, maar moeten nog verder ontwikkeld en/of gevalideerd worden. Het is van cruciaal belang dat dit gebeurt met modelstoffen die representatief zijn voor de microverontreinigingen die in bronnen voor drinkwater worden aangetroffen. Het samenstellen van een lijst met representatieve stoffen is echter een uitdaging. Er zijn uitgebreide lijsten bekend met verontreinigingen die voorkomen in bronnen van drinkwater, maar voor het huidige project was het van belang dat de range van alle voorkomende alle stoffeigenschappen zo goed mogelijk werd vertegenwoordigd, om modellen te krijgen die ook over die hele range kunnen worden toegepast. Daarom is in het in dit rapport beschreven onderzoek een objectieve methode ontwikkeld voor stofselectie. Overigens kunnen de kosten voor validatie in sommige gevallen worden beperkt door geen gebruik te maken van nieuwe experimentele data, maar van in de literatuur of eerder (KWR) onderzoek gepubliceerde metingen.

Bij deze methode zijn we uitgegaan van een lijst met 2400 stoffen, waarvan wiskundige descriptoren en vingerafdrukken bekend zijn. Op basis van “expert judgement” (in dit geval specifieke kennis van organische chemie en zuiveringsprocessen bij Roberta Hofman en Emile Cornelissen) zijn hieruit ongeveer 1400 parameters gekozen, die invloed kunnen hebben op het gedrag van een stof in adsorptie-, membraan- of oxidatieprocessen (de drie processen waarop het modelleringsonderzoek voor de drinkwaterbereiding zich zal richten, omdat deze in de belangstelling staan van de drinkwaterbedrijven). Vervolgens is onderzocht in welke range deze parameters voorkomen onder microverontreinigingen die de drinkwaterbedrijven rapporteren in hun bronnen. Daarna is met behulp van wiskundige technieken vastgesteld dat bij het toepassen van 30 clusters van stoffeigenschappen deze ranges zo goed mogelijk worden vertegenwoordigd, en is uit elk cluster één stof uitgekozen als representatief. Hierbij is rekening gehouden met het voorkomen van deze stoffen in bronnen voor drinkwater, en mogelijkheden om ze te analyseren. Verder is ervoor gezorgd dat de eigenschappen van de 30 uitgekozen stoffen zo min mogelijk overlap met elkaar vertonen.

Op deze manier is een lijst verkregen met 30 componenten, waarvan de eigenschappen representatief zijn voor de hele range aan relevante stoffeigenschappen voor de waterzuivering. Vervolgens is een analysemethode ontwikkeld, waarmee de meeste van deze stoffen in gelijktijdig gemeten kunnen worden, zodat de analysekosten beperkt kunnen worden.

De hier samengestelde lijst van representatieve modelstoffen zal worden gebruikt voor de validatie van modellen voor drinkwaterzuivering met behulp van filtratie over granulaire actieve kool, membraanfiltratie, en oxidatieprocessen gebaseerd op ozon.

Inhoud

Rapport	0
<i>Managementsamenvatting</i>	1
Bedrijfsparagraaf	7
Samenvatting	8
Inhoud10	
1 Algemene introductie	11
2 Introductie selectiemethode	14
3 Aanpak	15
4 Selectielijst	21
5 Conclusies	22
6 Referenties	23
I Data gebruikt voor stofselectie	24

1 Algemene introductie

Bronnen voor drinkwater bevatten steeds hogere concentraties (organische) microverontreinigingen, en bovendien neemt het aantal stoffen, dat wordt aangetroffen, alomarm toe. Een aantal jaar geleden werd al becijferd dat het om een jaarlijkse vracht van 300 miljoen ton zou gaan (Schwarzenbach, Escher et al. 2006). De verwachting is dat dit inmiddels is toegenomen: alleen al voor Nederland is een jaarlijkse lozing van minstens 140 ton per jaar becijferd (<https://www.rijksoverheid.nl/onderwerpen/geneesmiddelen/medicijnresten-in-water>). De afgelopen jaren is er wereldwijd uitgebreid onderzoek gedaan naar diverse technologieën om dergelijke verontreinigingen uit bronnen voor drinkwater te verwijderen, en ook in Nederland, in het kader van het Bedrijfstakonderzoek, zijn diverse projecten uitgevoerd. Technologieën die in dit kader zijn onderzocht, zijn bijvoorbeeld filtratie over granulaire actieve kool, toevoeging van poederkool, geavanceerde oxidatie (UV/H₂O₂, O₃, UV/O₃, O₃/H₂O₂), ultrafiltratie en omgekeerde osmose. Van tientallen stoffen is inmiddels bekend hoe ze zich in dergelijke processen gedragen. Voor nieuwe c.q. onbekende stoffen is het soms mogelijk om op basis van moleculaire structuurvergelijkingen en ervaring te voorspellen welke verwijderingstechniek effectief zou kunnen zijn, maar in veel gevallen zijn toch experimenten nodig om echt uitsluitsel te kunnen geven.

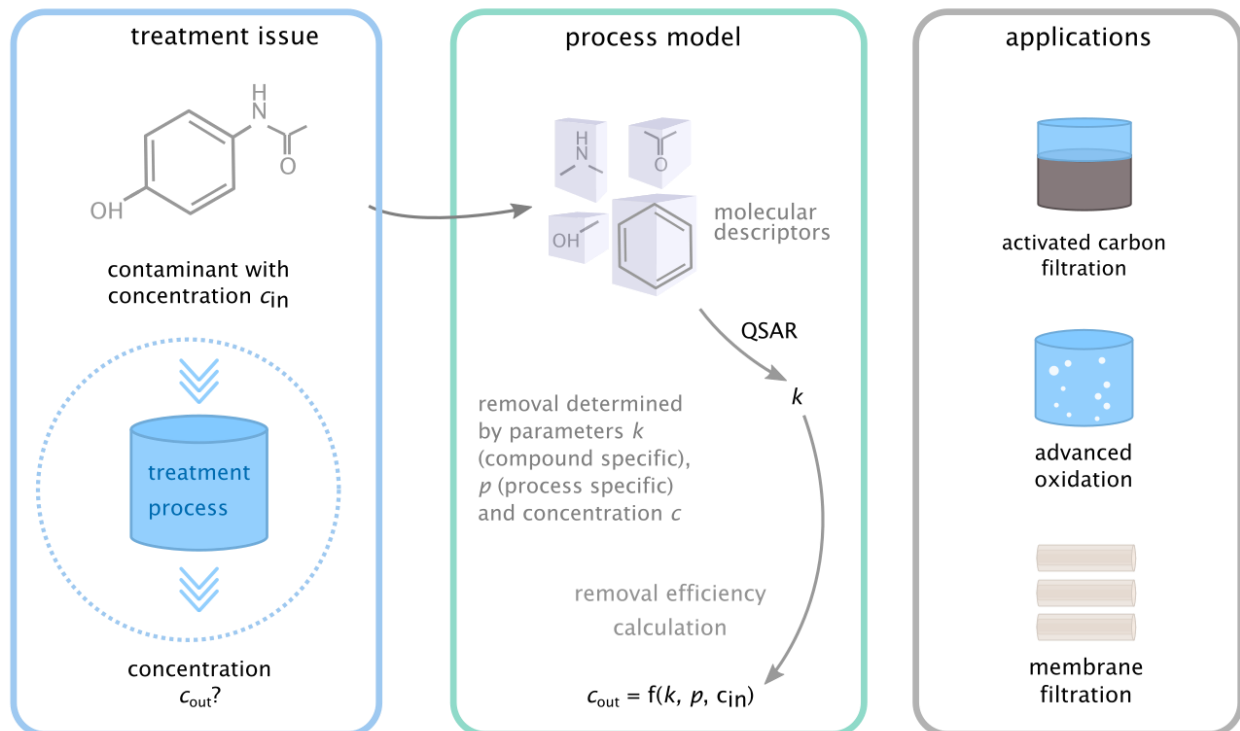
Dergelijke experimenten zijn erg duur, en het zou dus zeker een voordeel zijn als betrouwbaardere voorspellingen gedaan zouden kunnen worden. In dit verband lijken QSPRs een oplossing te kunnen bieden. “QSPR” staat voor “Quantitative Structure Property Relationship”. Het betekent dat er gezocht wordt naar een statistisch verband tussen de moleculaire structuur van een stof en zijn gedrag in bepaalde situaties. Die moleculaire structuur wordt wiskundig beschreven aan de hand van diverse parameters, waaronder afmetingen, aanwezigheid van functionele groepen, molmassa, enzovoort. Dergelijke verbanden kunnen alleen op een betrouwbare manier worden verkregen, als ze gebaseerd worden op een groot aantal experimentele gegevens. Daarnaast is een onafhankelijke set gegevens nodig om te controleren of een ontwikkelde QSPR ook inderdaad voldoet voor andere situaties of stoffen. Dit betekent dat er een grote set goed gedocumenteerde data nodig is, om een betrouwbare QSPR te kunnen ontwikkelen, en goede voorspellingen te kunnen doen.

Bij KWR is eerder al onderzoek gedaan naar de ontwikkeling van QSPRs voor nanofiltratie en geavanceerde oxidatie ((Verliefde, Cornelissen et al. 2005, Verliefde, Heijman et al. 2006, Heijman, Verliefde et al. 2007, De Ridder, McConville et al. 2009, Verliefde, Cornelissen et al. 2009, Yangali-Quintanilla, Verliefde et al. 2009, de Ridder, Villacorte et al. 2010, De Ridder, Verliefde et al. 2011; Wols and Vries 2012, Vries, Wols et al. 2013, Vries, Wols et al. 2017). Er is echter nog meer onderzoek nodig, om betrouwbare voorspellingen te kunnen doen voor allerlei andere waterbehandelingstechnieken. In dit onderzoek is gekeken of de modellen voor membraanfiltratie kunnen worden uitgebreid met andere typen membranen zoals omgekeerde osmose membranen. Hiermee zouden de modellen zowel voor traditioneel spiraal gewonden omgekeerde osmose membranen ook voor LBL (“layer by layer”) holle vezel- nanofiltratie membranen kunnen worden toegepast.

Adsorptieparameters van organische microverontreinigingen op actieve kool zijn in een eerder onderzoek gemodelleerd, maar hierbij kon nog geen rekening worden gehouden met de concurrentie die in de praktijk altijd optreedt door Natuurlijk Organisch Materiaal (NOM) (Hofman-Caris, Korevaar et al. 2018). In het huidige project is gepoogd die effecten wel mee te nemen. Details hiervan worden in deelrapport 2 van deze serie beschreven.

Wereldwijd worden er verschillende geavanceerde oxidatieprocessen toegepast in de drinkwaterzuivering. In Nederland wordt wel UV/H₂O₂ toegepast, maar in het buitenland zijn op ozon gebaseerde processen gangbaarder. Fotolyse door UV-straling en oxidatie door hydroxylradicalen, die ontstaan door fotolyse van H₂O₂, is gemodelleerd. Door tevens reacties die plaatsvinden onder invloed van ozon te modelleren, is het ook mogelijk geavanceerde oxidatieprocessen als O₃/H₂O₂ en O₃/UV te modelleren. Ook dit was onderdeel van het huidige onderzoeksproject.

Een overzicht van de toegepaste modelbenadering die voor alle behandelingsmethoden is gebruikt, is gegeven in Figuur 1.



Figuur 1: modelbenadering voor alle waterbehandelingsprocessen

Het idee is dus om met behulp van moleculaire software de stoffeigenschappen van moleculen te bepalen (de zogenaamde "descriptor"). Een statistisch model legt vervolgens verband tussen deze stoffeigenschappen en een eigenschap of reactieconstante in een proces. Voor een nieuwe stof kunnen dan de stoffeigenschappen berekend worden, en daarmee kan de eigenschap of reactieconstante voor die stof worden voorspeld. Met behulp van het procesmodel leidt dit tot een voorspelde verwijdering van de stof in een zuiveringsproces.

Om goede QSPRs te kunnen ontwikkelen, is het nodig een representatieve stoffenset te gebruiken, waarin alle eigenschappen worden vertegenwoordigd die van invloed zijn op het gedrag van stoffen in een zuivering (zoals molmassa, molecuulvolume, lading en ladingsverdeling, aanwezigheid van functionele groepen enz.). Een uitgebreid overzicht van alle parameters is weergegeven in Tabel 3 in Bijlage I. Daarom was de eerste taak binnen dit project het vaststellen van een dergelijke representatieve lijst. Dit rapport beschrijft op welke manier we geprobeerd hebben om objectief een representatieve lijst van microverontreinigingen samen te stellen.

Leeswijzer

Dit rapport is het eerste in een serie van vier rapporten over voorspelling en validatie van verwijderingsprocessen voor microverontreinigingen in water. In dit rapport wordt een overzicht gegeven van de werkwijze die is gehanteerd om een selectie te maken van representatieve stoffen, en van de uiteindelijke lijst die in dit onderzoek is gebruikt.

In deel twee van deze serie worden ozonatieprocessen ontrafeld, en gemodelleerd aan de hand van in de literatuur gevonden reactieconstanten. Voorspellingen worden vergeleken met data die op labschaal zijn verkregen. In deel drie zullen de resultaten worden besproken van het modelleren van membraanprocessen. In deel vier komen adsorptieprocessen (actieve kool) aan bod.

Het project is begeleid door een groep drinkwaterexperts, die bestond uit Jantinus Bruins (WLN), Liesbeth Verdickt (De Watergroep), Menno van Leenen (Oasen), René van der Aa (Waternet), Rinnert Schurer (Evides) en David de Ridder (Evides).

2 Introductie selectiemethode

Voor BTO projecten worden regelmatig experimenten met organische microverontreinigingen (OMV) uitgevoerd. Daarom is een lijst samengesteld met stoffen die representatief zijn voor de belangrijkste zuiveringsprocessen die op dit moment worden toegepast: membraanfiltratie, actieve koolfiltratie en (geavanceerde) oxidatie (biologische processen zijn in dit project nog niet meegenomen). Voor de kwalificatie “representatief” is gekeken naar een aantal factoren:

- De geselecteerde stoffen moeten eigenschappen hebben die een rol spelen in de effectiviteit van een bepaalde zuiveringsstap. Denk hierbij bijvoorbeeld aan de molmassa en het aantal aromatische groepen: factoren die een rol spelen in de mate van adsorptie op actieve kool, de verwijdering met behulp van bepaalde membranen, of de oxideerbaarheid van een verbinding.
- De geselecteerde stoffen, of stoffen die hier chemisch sterk op lijken, worden aangetroffen in bronnen voor drinkwater.

De moleculaire structuur van stoffen kan op verschillende manieren worden beschreven. “Fingerprints” zijn structurelementen die erin voorkomen, zoals een zuurgroep, een ester, een amine of aromaten. Daarnaast zijn er nog “descriptorren”: numerieke waarden van stoffen, zoals molmassa, aantal halogenen enzovoort. De gebruikte fingerprints en descriptorren zijn weergegeven in Tabel 3 en Tabel 4 in Bijlage I. Op basis van wiskundig beschreven stoffeigenschappen en vingerafdrukken¹ van ~2400 stoffen in de AquaPriori database is gezocht naar een lijst van ongeveer 30 stoffen die representatief is voor de range aan stoffeigenschappen die microverontreinigingen in water laten zien, en die een rol spelen in de verwijderingsmechanismen van bovengenoemde drie zuiveringsprocessen. Hierbij is rekening gehouden met de voor de bedrijven relevante stoffen. Bovendien is gekeken naar bestaande analysemethoden, omdat het voor het onderzoek eenvoudiger en goedkoper is als zo min mogelijk verschillende analysemethoden moeten worden toegepast. De stoffen die door de bedrijven zijn genoemd en/of die voorkomen in een analysemethode zijn weergegeven in Bijlage I.

In hoofdstuk 3 beschrijven we de aanpak om tot deze geselecteerde stoffenlijst te komen. Vervolgens heeft het KWR lab een analysemethode ontwikkeld waarmee deze lijst van stoffen in één keer geanalyseerd kan worden.

¹ Onder “vingerafdruk” worden structurelementen als een zuurgroep, een ester, ether, aldehyde enzovoort in een molecuul verstaan.

3 Aanpak

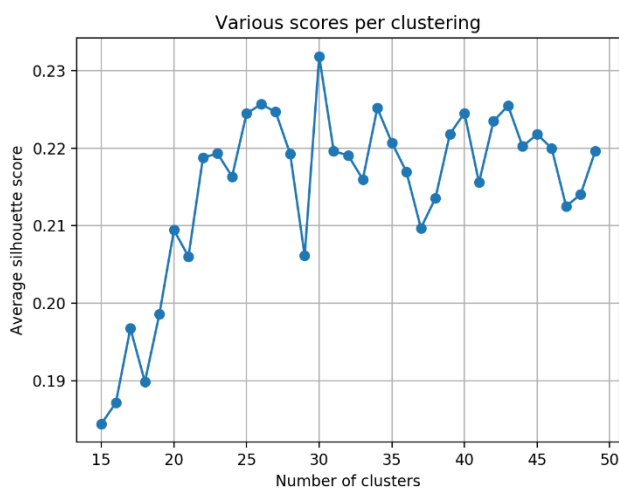
Bij de keuze van stoffen voor de stoffenlijst voor modellering is gebruik gemaakt van bestaande lijsten: analyselijsten (lijsten van stoffen die met bestaande analysemethoden geanalyseerd kunnen worden), en bedrijvenlijsten (lijsten van stoffen die door de drinkwaterbedrijven genoemd worden). Deze zijn samengevat in Tabel 2, Bijlage I.

Voor de samenstelling van de stoffenlijst zijn de volgende stappen gevolgd:

1). Berekening descriptoren: Op basis van expertkennis binnen KWR zijn de verschillende wiskundige descriptoren waarover we beschikken beoordeeld, en de relevante descriptoren per zuiveringsproces geselecteerd (zie Tabel 3 en Tabel 4 in Bijlage I). Voor membraanprocessen zijn vooral grootte, lading en polariteit van belang. Voor adsorptie spelen deze parameters ook een rol, maar daarnaast ook zaken als de aanwezigheid van aromatische groepen, lengte en vertakking van koolstofketens en dergelijk, en voor oxidatie de aanwezigheid van structurelementen die wel of niet geoxideerd kunnen worden. Dit leidde tot een lijst met ongeveer 1400 descriptoren en fingerprints, die vervolgens tot één lijst zijn gecombineerd. Deze lijst met alle descriptoren staat in Bijlage I (Tabel 3 en Tabel 4).

2). Clustering: Op de geselecteerde descriptoren is eerst een principal component analysis (PCA) uitgevoerd, waarbij 20 'principal components' zijn geselecteerd. Dit zijn de belangrijkste descriptoren die 53% van de variabiliteit in de data van alle stoffen verklaren. Dit is nodig, omdat de clustering anders niet goed werkt (dan is er te weinig onderscheid tussen de clusters).

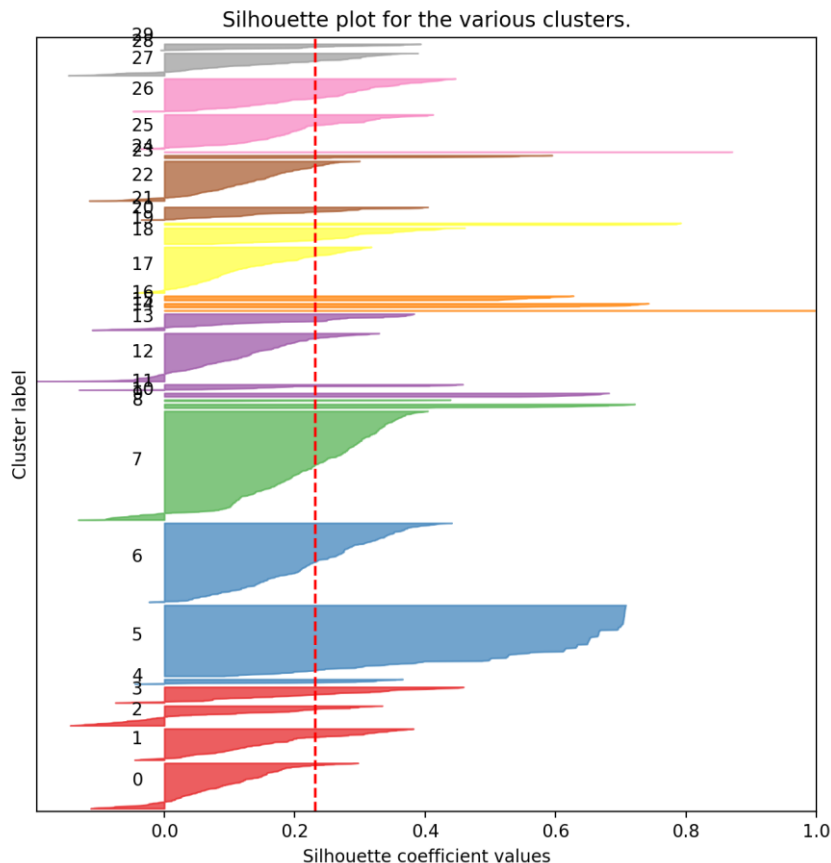
Een clusteranalyse is uitgevoerd op alle stoffen, waarbij met 15 tot 50 clusters is gevarieerd. Het optimale aantal clusters is vervolgens gekozen op basis van de gemiddelde silhouette-score² (deze score geeft aan hoe goed een stof past binnen een cluster). De gemiddelde silhouette score per aantal gekozen clusters is weergegeven in Figuur 2.



Figuur 2: Gemiddelde silhouette score per aantal gekozen clusters

² Silhouette analyse wordt gebruikt om de afstand tussen clusters in te schatten. Silhouette scores van +1 geven aan dat een stof sterk verschilt met andere clusters. Waardes van 0 of lager geven aan dat een stof mogelijk aan het verkeerde cluster toegekend is.

Hieruit blijkt dat tussen 15 en 25 clusters de gemiddelde silhouette score toeneemt, en bij 30 clusters een maximum bereikt. De silhouette grafiek voor deze 30 clusters is weergegeven in Figuur 3. Hierin wordt voor alle 2400 stoffen uit de database de silhouette score per cluster weergegeven. Als de silhouette score kleiner is dan 0 wordt die stof niet goed vertegenwoordigd door het desbetreffende cluster. Dit blijkt het geval te zijn voor een klein aantal stoffen.



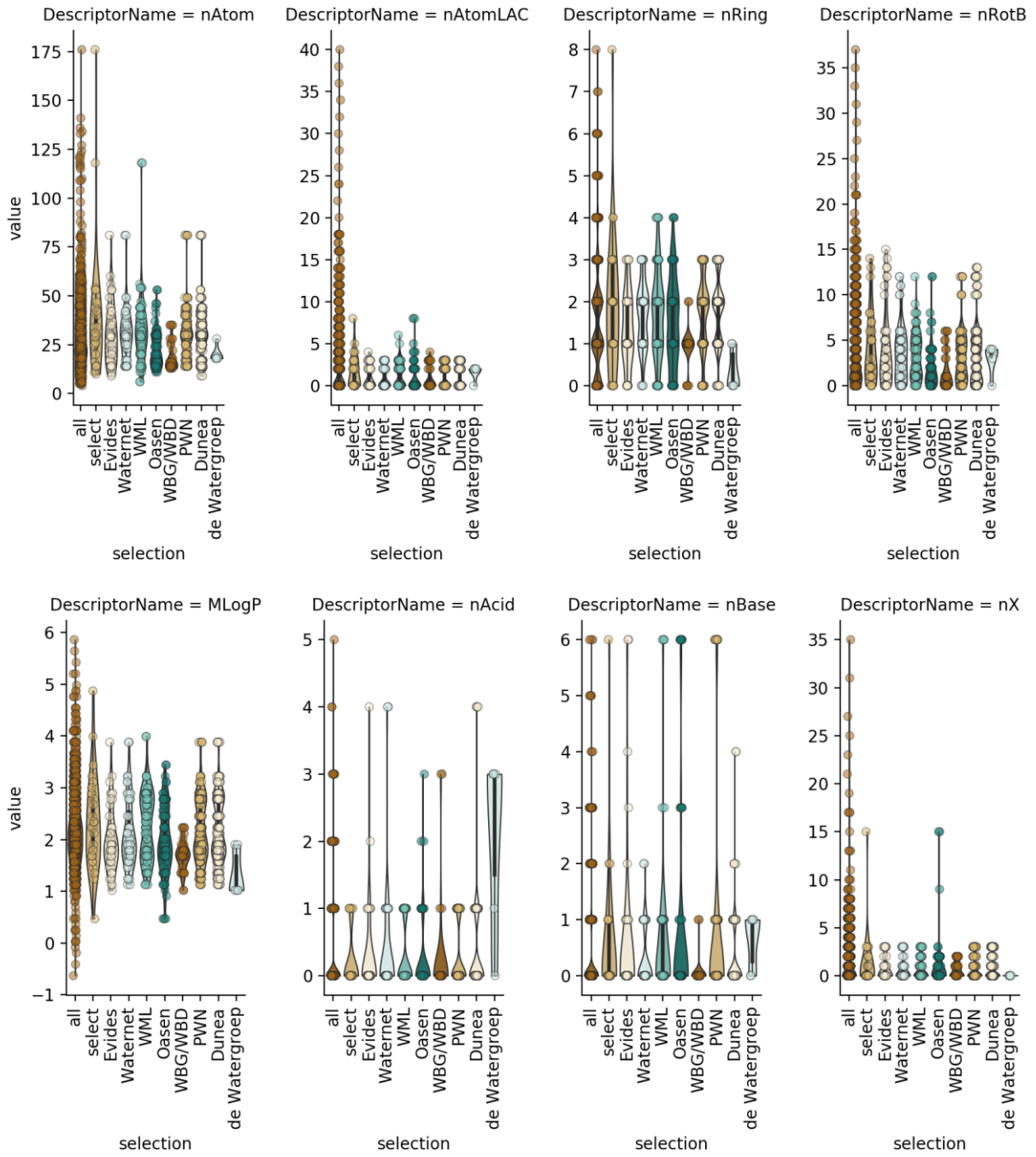
Figuur 3: Silhouette grafiek van 30 clusters. De silhouette score geeft aan hoe goed een stof in een bepaald cluster past: hoe hoger hoe beter. Als een stof een silhouette score < 0 heeft, dan wordt deze niet goed vertegenwoordigd door het desbetreffende cluster.

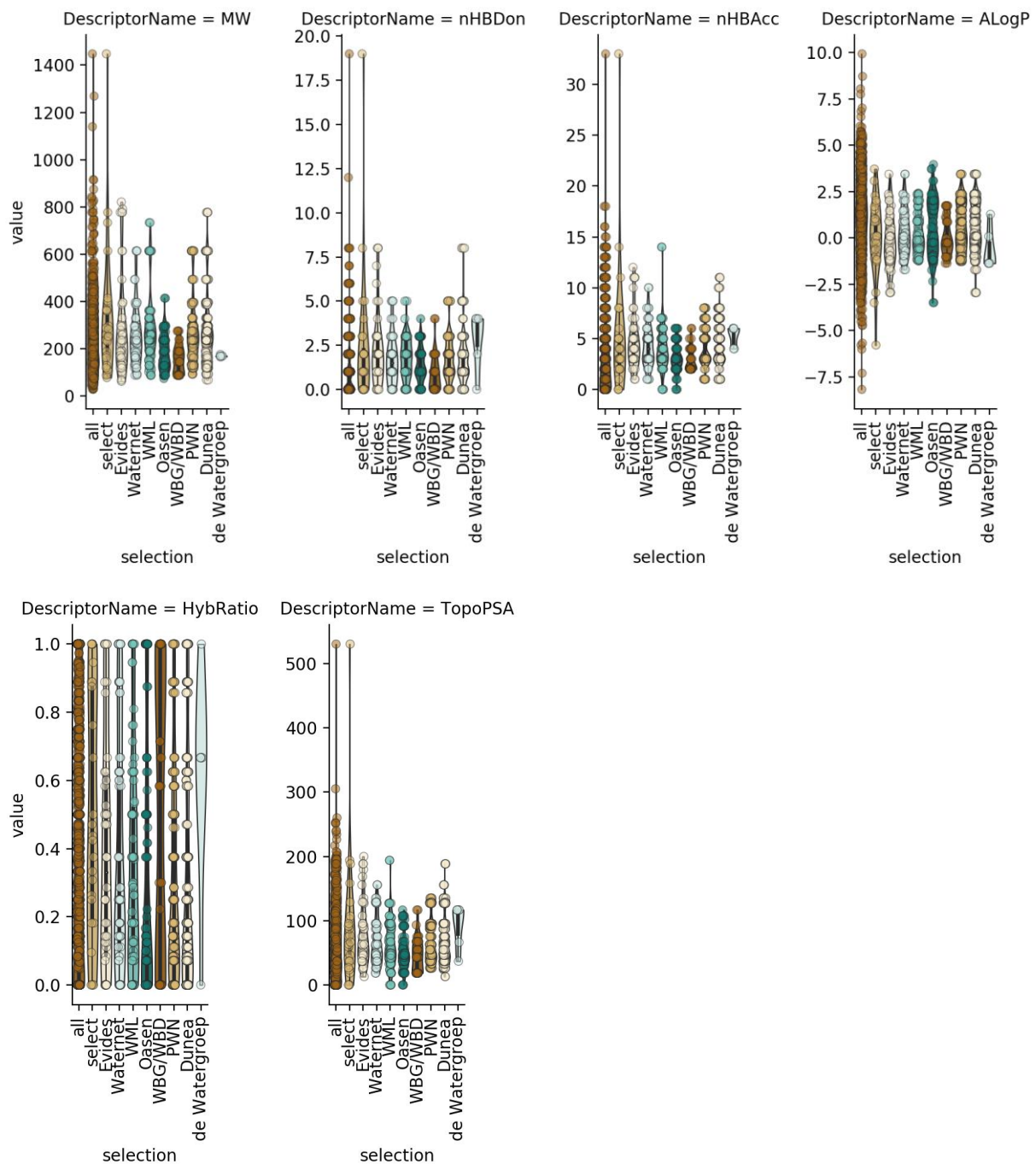
3) Selectie van stoffen uit ieder cluster: Van ieder cluster is één stof gekozen volgens de volgende criteria:

- Van de stoffen die momenteel in de analyselijsten (lijsten van componenten in bestaande analysemethoden) voorkomen, is per cluster een stof gekozen die het meeste genoemd wordt door de verschillende bedrijven. Dit betreft 17 clusters.
- Als er in een cluster geen stoffen zijn die in de analyselijsten voorkomen, is ook de stof gekozen die het meeste genoemd wordt door de verschillende bedrijven. Dit betreft 6 clusters. Bij gelijke aantallen is de stof gekozen met de hoogste silhouette-score.
- Als er in een cluster wel stoffen zijn die in de analyselijsten voorkomen, maar niet in de bedrijvenlijsten (Tabel 2 in Bijlage 1), wordt een stof geselecteerd die qua naam het meest bekend voorkomt bij de auteurs en de begeleidingsgroep en een hoge silhouette-score heeft. Dit betreft 1 cluster.
- Als er in een cluster geen stoffen zijn die in de analyselijsten voorkomen en ook niet in de lijsten die geleverd zijn door de bedrijven, wordt handmatig een stof geselecteerd die qua naam het meest bekend voorkomt. Dit betreft 6 clusters.

- Daarnaast is gekeken naar verkrijgbaarheid (en de kosten hiervan) voor de stoffen die nog niet in een analysemethode zitten. Op basis hiervan zijn 2 stoffen nog vervangen door een andere stof uit hetzelfde cluster (cluster 19 en 24).

De geselecteerde stoffen worden geëvalueerd met een afstandsmetrik bepaald op basis van een aantal bekende descriptorwaarden en een vioolplot (Figuur 4) die de verdelingen van descriptorwaarden voor de stoffenlijsten toont.



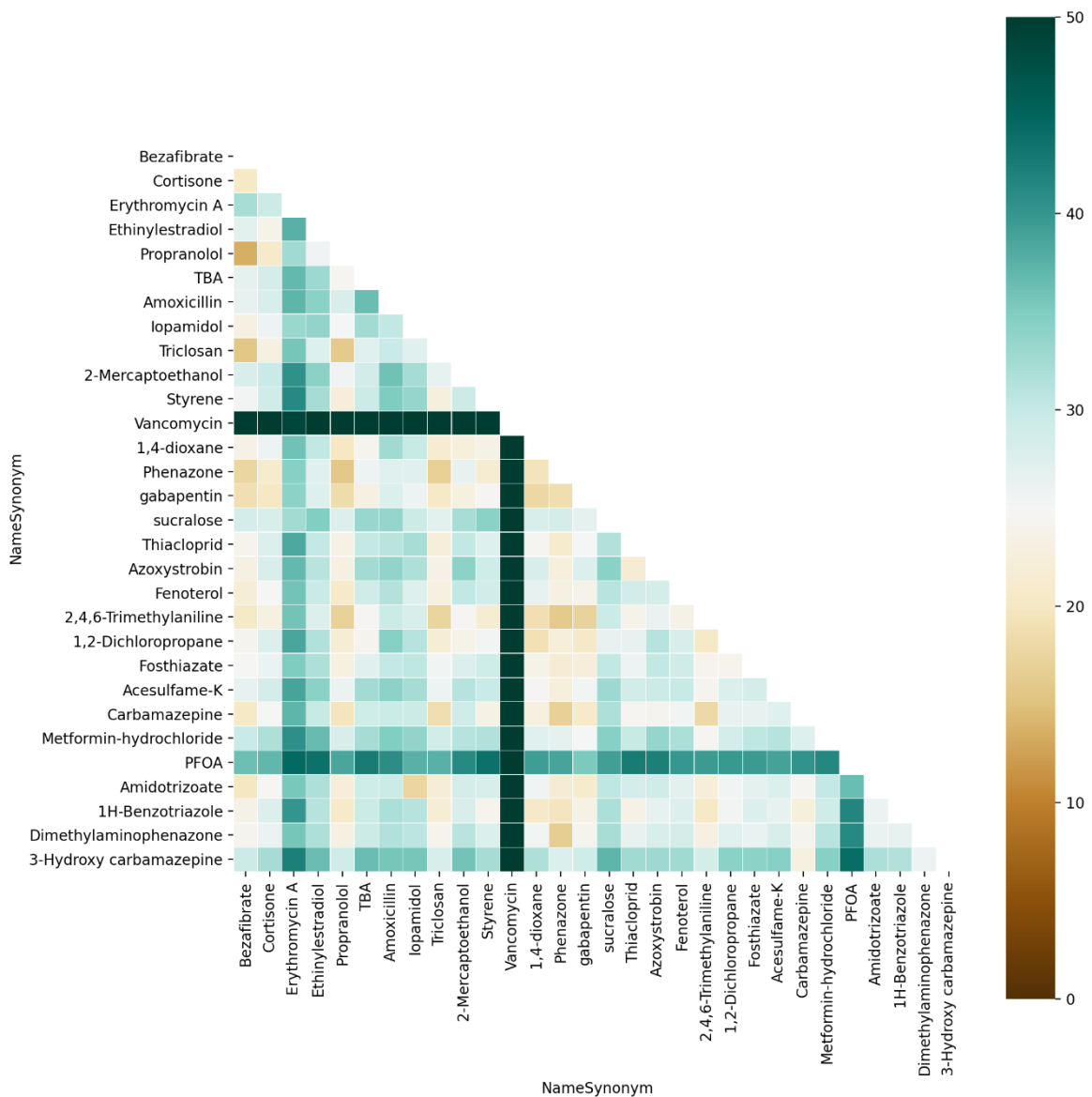


Figuur 4 Verdeling van verschillende descriptorwaarden in vioolgrafiek. ‘all’ staat voor alle 2400 stoffen, ‘select’ is de opgestelde stoffenlijst. Deze lijsten zijn vergeleken met de lijsten van de verschillende bedrijven.

De “vioolplots” in Figuur 4 laten zien hoeveel stoffen uit een set een bepaalde waarde van een descriptor hebben. Naarmate de plot breder is, vertonen meer stoffen uit de selectie die waarde voor de betreffende eigenschap. Een “buik” in de plot betekent dan dat relatief veel stoffen die waarde vertonen. Hoe minder stoffen een bepaalde waarde vertonen, hoe smaller de plot wordt. In sommige gevallen is er bijvoorbeeld één stof die een veel hogere of lagere waarde vertoont, wat resulteert in een lijn naar die waarde toe. We hebben geprobeerd een selectie te maken die de range aan (relevante) eigenschappen die voorkomt (weergegeven door de plot voor “all”) zo goed mogelijk benadert. Hierbij is er vooral op gelet dat de eigenschappen (descriptor en fingerprints) van door de bedrijven gerapporteerde microverontreinigingen optimaal benadert wordt. Verder blijkt uit deze grafieken dat de

geselecteerde stoffen samen een breder palet aan eigenschappen bestrijken dan de lijstjes van de individuele bedrijven.

Vervolgens is een “afstand” grafiek gemaakt, die laat zien in hoeverre stoffen op elkaar lijken. Er is voor gezorgd dat de stoffen in onze selectie zo min mogelijk op elkaar lijken, om een zo groot mogelijke range van eigenschappen te kunnen bestrijken. Dit is weergegeven in Figuur 5. Naarmate de eigenschappen van twee stoffen meer van elkaar verschillen is de waarde in deze grafiek hoger en de kleur dus donkerder. Dezelfde stoffen horizontaal en verticaal scoren een wit blokje in dit overzicht, en stoffencombinaties met een lichtgekleurd blokje vertonen veel overeenkomsten. Stoffencombinaties met een donkergroen blokje verschillen het meest van elkaar. Met de “afstand” grafiek kunnen stoffen die erg op elkaar lijken ook nog geschrapt worden, maar dat bleek in dit geval niet nodig.



Figuur 5: Afstand tussen stoffen, hoe hoger de waarde, hoe groter het verschil is tussen de stoffen.

Op basis van deze data is een keuze gemaakt voor een aantal stoffen.

4 Selectielijst

De (via de in hoofdstuk 3 beschreven methode) geselecteerde lijst met stoffen is weergegeven in Tabel 1. Het bleek niet mogelijk te zijn drie van deze stoffen met de beoogde methode te analyseren. Die stoffen zijn dan uit de lijst gehaald. Een compleet overzicht is gegeven in Tabel 1

Tabel 1: overzicht van geselecteerde stoffen, met het cluster dat zij vertegenwoordigen of ze al in een bestaande analysemethode voorkwamen. # bedrijven = het aantal drinkwaterbedrijven dat de aanwezigheid van deze verontreiniging gerapporteerd heeft.

NameSynonym	CAS	Cluster	In een bestaande analysemethode	# Bedrijven	analysemethode
Carbamazepine	298-46-4	0	V	8	1
Gabapentin	60142-96-3	1	V	6	1
Amidotrizoate	117-96-4	2	V	6	1
TBA	1112-67-0	3		1	1
Iopamidol	60166-93-0	4		3	1
Triclosan	3380-34-5	5		1	geen
1,4-dioxane	28347-88-8	6	V	4	2
Phenazone	60-80-0	7	V	2	1
PFOA	335-67-1	8		1	1
3-Hydroxy carbamazepine	68011-67-6	9		1	1
Ethinylestradiol	57-63-6	10		1	4
Azoxystrobin	131860-33-8	11		0	1
Bezafibrate	41859-67-0	12	V	2	1
Cortisone	53-06-5	13	V	1	1
Metformin	657-24-9	14	V	5	5
Erythromycin A	114-07-8	15	V	1	1
Amoxicillin	26787-78-0	16		0	3
Benzotriazole	95-14-7	17	V	7	1
Styrene	79637-11-9	18	V	2	2
2-Mercaptoethanol	60-24-2	19		0	geen
Thiacloprid	111988-49-9	20	V	1	1
Fenoterol	13392-18-2	21		0	1
Propranolol	525-66-6	22	V	7	1
Dimethylaminophenazone	58-15-1	23	V	0	1
2,4,6-Trimethylaniline	88-05-1	24		0	1
Acesulfame	33665-90-6	25	V	5	1
Fosthiazate	98886-44-3	26	V	1	1
1,2-Dichloropropane	78-87-5	27	V	3	2
Sucralose	56038-13-2	28	V	5	1
Vancomycin	1404-90-6	29		0	geen

1 = "Lijst Roberta" C18 methode

- 2 = vluchtigen methode
- 3 = antibiotica methode, stabiliteit van de standaarden nog onbekend
- 4 = hormonen methode
- 5 = HILIC methode
- Geen = deze stof hebben we niet in een methode zitten

Stoffen die niet door bedrijven gerapporteerd zijn kunnen toch relevant zijn, omdat zij bepaalde eigenschappen combineren. Vervolgens is er door het laboratorium van KWR een analysemethode ontwikkeld waarmee de meeste van de hier boven genoemde stoffen in één keer geanalyseerd kunnen worden (methode 1). Hierdoor is voor analyse van de hierboven genoemde set aan stoffen slechts een beperkt aantal (hooguit vijf) methodes nodig (vluchtige stoffen worden apart gemeten, maar die zijn niet in alle series experimenten meegenomen). Op deze manier kunnen de kosten voor verder onderzoek in het kader van dit project relatief laag gehouden worden.

5 Conclusies

In dit deel van het project “Voorspelling en validatie van verwijdering van organische microverontreinigingen” is een methode ontwikkeld waarmee op een objectieve manier een selectie van modelcomponenten gemaakt kan worden. Op basis van de descriptors en fingerprints die waarschijnlijk een rol spelen in een bepaald zuiveringsproces is via wiskundige methodes een selectie gemaakt van 30 organische stoffen, die zo goed mogelijk de hele range aan relevante stofeigenschappen van organische microverontreinigingen in water vertegenwoordigen. Vervolgens is een analysemethode ontwikkeld, waarmee deze combinatie van stoffen tegelijkertijd geanalyseerd kan worden.

De aldus samengestelde lijst aan microverontreinigingen is in het vervolgonderzoek (modellering van adsorptie-, membraan- en oxidatieprocessen) gebruikt. De resultaten van dit onderzoek zullen in de delen 2, 3 en 4 van deze reeks rapporten in detail worden besproken.

6 Referenties

De Ridder, D. J., M. McConville, A. R. D. Verliefde, L. T. J. Van Der Aa, S. G. J. Heijman, J. Q. J. C. Verberk, L. C. Rietveld and J. C. Van Dijk (2009). "Development of a predictive model to determine micropollutant removal using granular activated carbon." *Drinking Water Engineering and Science* 2(2): 57-62.

De Ridder, D. J., A. R. D. Verliefde, S. G. J. Heijman, J. Q. J. C. Verberk, L. C. Rietveld, L. T. J. Van Der Aa, G. L. Amy and J. C. Van Dijk (2011). "Influence of natural organic matter on equilibrium adsorption of neutral and charged pharmaceuticals onto activated carbon." *Water Science and Technology* 63(3): 416-423.

de Ridder, D. J., L. Villacorte, A. R. D. Verliefde, J. Q. J. C. Verberk, S. G. J. Heijman, G. L. Amy and J. C. van Dijk (2010). "Modeling equilibrium adsorption of organic micropollutants onto activated carbon." *Water Research* 44(10): 3077-3086.

Heijman, S. G. J., A. R. D. Verliefde, E. R. Cornelissen, G. Amy and J. C. Van Dijk (2007). Influence of natural organic matter (NOM) fouling on the removal of pharmaceuticals by nanofiltration and activated carbon filtration. *Water Science and Technology: Water Supply*. 7: 17-23.

Hofman-Caris, C. H. M., M. Korevaar, W. G. Siegers, D. Vries and B. A. Wols (2018). Doorbraakcurves voor het vaststellen van reactivatiecriteria voor actieve kool, BTO 2018.087, KWR Water Research Institute, Nieuwegein.

Schwarzenbach, R. P., B. I. Escher, K. Fenner, T. B. Hofstetter, C. A. Johnson, U. Von Gunten and B. Wehrli (2006). "The challenge of micropollutants in aquatic systems." *Science* 313(5790): 1072-1077.

Vries, D., B. A. Wols and P. De Voogt (2013). "Removal efficiency calculated beforehand: QSAR enabled predictions for nanofiltration and advanced oxidation." *Water Science and Technology: Water Supply* 13(6): 1425-1436.

Vries, D., B. A. Wols, M. Korevaar and E. Vonk (2017). AquaPriori: a priori het verwijderingsrendement bepalen, KWR Watercycle Research Institute, Nieuwegein.

Wols, B. A. and D. Vries (2012). "On a QSAR approach for the prediction of priority compound degradation by water treatment processes." *Water Science and Technology* 66(7): 1446-1453.

Verliefde, A., E. Cornelissen, G. Amy, B. Van Der Bruggen and H. Van Dijk (2005). A QSAR/QSPR-methodology for rejection of trace organic pollutants in nanofiltration. 2005 Water Quality Technology Conference Proceedings, WQTC 2005.

Verliefde, A., B. Heijman, G. Amy, B. Van Der Bruggen and J. C. Van Dijk (2006). Influence of hydrophobicity and charge on removal of trace organic pollutants with high pressure membranes. American Water Works Association - Water Quality Technology Conference and Exposition 2006: Taking Water Quality to New Heights.

Verliefde, A. R. D., E. R. Cornelissen, S. G. J. Heijman, I. Petrinic, T. Luxbacher, G. L. Amy, B. Van der Bruggen and J. C. van Dijk (2009). "Influence of membrane fouling by (pretreated) surface water on rejection of pharmaceutically active compounds (PhACs) by nanofiltration membranes." *Journal of Membrane Science* 330(1-2): 90-103.

Yangali-Quintanilla, V., A. Verliefde, T. U. Kim, A. Sadmani, M. Kennedy and G. Amy (2009). "Artificial neural network models based on QSAR for predicting rejection of neutral organic compounds by polyamide nanofiltration and reverse osmosis membranes." *Journal of Membrane Science* 342(1-2): 251-262.

4-t-Octylphenol (140-66-9)			v						
5,6-dimethyl-2H-benzotriazole (4184-79-6)									v
5-chloro-2H-benzotriazole (94-97-3)									v
5-hydroxyquinoline (578-67-6)			v						
5-Methyl-1H-benzotriazole (136-85-6)	v			v	v	v			v
6-hydroxyquinoline (580-16-5)			v						
Acesulfame-K (33665-90-6)	v			v	v	v		v	v
Acetamidrid (160430-64-8)									v
Acetylsulfamethoxazole (21312-10-7)									v
Acrylonitrile (63908-52-1)									v
Aldicarb (116-06-3)								v	
allylchloride (107-05-1)								v	v
Amidosulfuron (120923-37-7)									v
Amidotrizoate (117-96-4)	v	v		v	v	v	v		v
Aminophenazone (58-15-1)									v
AMPA (1066-51-9)				v			v	v	
Anhydro erythromycin A (23893-13-2)									v
Aniline (62-53-3)	v			v	v	v		v	
Atenolol (60966-51-0)		v					v	v	v
Atrazine (1912-24-9)		v	v				v	v	
Azimsulfuron (120162-55-2)									v
barbital (57-44-3)	v			v	v	v			v
Bentazon (25057-89-0)		v					v	v	
Benzene (27271-55-2)							v	v	v
Bezafibrate (41859-67-0)		v					v		v
Bisoprolol (66722-44-9)		v						v	
Bisphenol A (27100-33-0)			v					v	
Bromacil (314-40-9)							v	v	
Bromobenzene (108-86-1)							v	v	v
Butalbital (77-26-9)									v
Butoxycarboxim (34681-23-7)								v	
Caffeine (71701-02-5)		v	v	v				v	
Candesartan (139481-59-7)									v
Carbamazepine (298-46-4)	v	v	v	v	v	v	v	v	v
Carbendazime (63278-70-6)	v			v	v	v	v	v	
Chloridazon desphenyl (6339-19-1)	v	v	v	v			v	v	v
Chlorobenzene (68411-45-0)							v	v	v
Chloroethene (75-01-4)								v	
Cis-1,2-dichloroethylene (156-59-2)		v					v	v	v
Clenbuterol (50499-60-0)		v							v
Clindamycin (18323-44-9)							v		v
Clofibric acid (882-09-7)		v					v	v	v
Cortisol (50-23-7)		v							v
Cortisone (53-06-5)		v							v
Cumene (51609-83-7)									v
cyclohexane (25012-93-5)								v	v
Cyclophosphamide (50-18-0)		v							v

MCPA (94-74-6)							v	v	
m-Dichlorobenzene (541-73-1)							v	v	v
Mecoprop (7085-19-0)							v	v	
mesitylene (108-67-8)							v		v
Mesotrione (104206-82-8)							v		v
Metconazole (125116-23-6)							v		v
Metformin-hydrochloride (1115-70-4)		v	v	v	v		v		v
Methyl tertiary-butyl ether (1634-04-4)							v	v	v
Methylene chloride (1605-72-7)							v	v	v
methylimino-sulfanylidene methane (556-61-6)									v
Metolachlor (55762-76-0)							v	v	
Metoprolol (37350-58-6)		v					v	v	v
Metronidazole (99616-64-5)		v							v
m-Xylene (108-38-3)							v	v	v
N,N-dibutyl nitrous amide (924-16-3)									v
N,N-diethyl-2-naphthalen-1-yloxypropanamide (15299-99-7)									v
N,N-diethylnitrous amide (55-18-5)									v
N,N-dipropyl nitrous amide (621-64-7)									v
Naphtalene (72931-45-4)							v	v	v
Naproxen (26159-31-9)		v					v	v	v
NDMA (62-75-9)			v						v
N-ethyl-N-methylnitrous amide (10595-95-6)									v
Niacin (123574-58-3)		v							v
Nicosulfuron (111991-09-4)							v	v	v
nitroguanidine (556-88-7)			v						
N-methyl-2-pyridinamine (4597-87-9)			v						
NMOR (59-89-2)									v
Norfluoxetine (83891-03-6)									v
o-Chlorotoluene (95-49-8)							v	v	v
O-Desmethyl Metoprolol (62572-94-5)									v
Oxasulfuron (144651-06-9)									v
Oxcarbamazepine (28721-07-5)									v
o-Xylene (95-47-6)							v	v	v
Oxypurinol (2465-59-0)									v
Paracetamol (8055-08-1)		v	v	v			v	v	v
Paroxetine (63952-24-9)		v							v
pCBA (74-11-3)		v							
p-Chlorotoluene (106-43-4)							v	v	v
Pentobarbital (76-74-4)									v
Pentoxifylline (6493-05-6)		v					v		v
Perfluorobutanesulfonic acid (375-73-5)			v						
p-fluorobenzoic acid (456-22-4)			v						
PFOA (335-67-1)			v						
Phenazone (60-80-0)		v					v		v
Phenobarbital (11097-06-6)	v			v	v	v			v
PHENYLUREA (64-10-8)			v						

Triclosan (3380-34-5)			v						
Triethyl phosphate (78-40-0)			v						
Triflusulfuron-methyl (126535-15-7)								v	v
Triglyme (70992-85-7)								v	
Trimethoprim (738-70-5)		v						v	v
Tri-N-buthylfosfaat (80094-39-9)				v					
Triphenylphosphine oxide (791-28-6)	v			v	v	v			
Uracil (66-22-8)			v						
Urotropin (24911-88-4)	v			v					v
Valsartan (137862-53-4)									v
Venlafaxine (93413-69-5)		v							v

Tabel 3: Lijst met descriptoren die relevant geacht worden voor adsorptie op actieve kool, membraanprocessen of oxidatieprocessen. De parameters zijn weergegeven in hun Engelse benaming, zoals ze in de literatuur en in modellen worden gebruikt. TRUE = relevant voor het betreffende proces; FALSE = niet relevant voor het betreffende proces.

Descriptor type	Descriptor	Actieve kool	membranen	Oxidatie
Acidic group count	nAcid	TRUE	TRUE	TRUE
ALOGP	ALogP, ALogP2, AMR	TRUE	TRUE	FALSE
APol	apol	TRUE	TRUE	TRUE
Aromatic atoms count	naAromAtom	TRUE	TRUE	TRUE
Aromatic bonds count	nAromBond	TRUE	TRUE	TRUE
Atom count	nAtom, nHeavyAtom, nH, nB, nC, nN, nO, nS, nP, nF, nCl, nBr, nI, nX	TRUE	TRUE	FALSE
Basic group count	nBase	TRUE	TRUE	TRUE
Bond count	nBonds, nBonds2, nBondsS, nBondsS2, nBondsS3, nBondsD, nBondsD2, nBondsT, nBondsQ, nBondsM	FALSE	TRUE	FALSE
Carbon types	C1SP1, C2SP1, C1SP2, C2SP2, C3SP2, C1SP3, C2SP3, C3SP3, C4SP3	FALSE	FALSE	TRUE
Chi chain	SCH-3, SCH-4, SCH-5, SCH-6, SCH-7, VCH-3, VCH-4, VCH-5, VCH-6, VCH-7	FALSE	TRUE	FALSE
Crippen logP and MR	CrippenLogP, CrippenMR	FALSE	TRUE	FALSE
Atom type electrotopological state	nHBd, nwHBd, nHBa, nwHBa, nHBint2, nHBint3, nHBint4, nHBint5, nHBint6, nHBint7, nHBint8, nHBint9, nHBint10, nHsOH, nHdNH, nHsSH, nHsNH2, nHssNH, nHaaNH, nHsNH3p, nHssNH2p, nHssNHp, nHtCH, nHdCH2, nHdsCH, nHaaCH, nHCHnX, nHCsats, nHCsatu, nHAvin, nHother, nHmisc, nsLi, nssBe, nssssBem, nsBH2, nssBH, nssB, nssssBm, nsCH3, ndCH2, nssCH2, ntCH,	FALSE	FALSE	TRUE

	<p>ndsCH, naaCH, nsssCH, nddC, ntsC, ndssC, naasC, naaaC, nssssC, nsNH3p, nsNH2, nssNH2p, ndNH, nssNH, naaNH, ntN, nsssNHp, ndsN, naaN, nsssN, nddsN, naasN, nssssNp, nsOH, ndO, nssO, naaO, naOm, nsOm, nsF, nsSiH3, nssSiH2, nssssSiH, nssssSi, nsPH2, nssPH, nsssP, ndsssP, nddsP, nsssssP, nsSH, ndS, nssS, naaS, ndssS, nddssS, nssssssS, nSm, nsCl, nsGeH3, nssGeH2, nsssGeH, nssssGe, nsAsH2, nssAsH, nsssAs, ndsssAs, nddsAs, nsssssAs, nsSeH, ndSe, nssSe, naaSe, ndssSe, nssssssSe, nddssSe, nsBr, nsSnH3, nssSnH2, nsssSnH, nssssSn, nsl, nsPbH3, nssPbH2, nsssPbH, nssssPb, SHBd, SwHBd, SHBa, SwHBa, SHBint2, SHBint3, SHBint4, SHBint5, SHBint6, SHBint7, SHBint8, SHBint9, SHBint10, SHsOH, SHdNH, SHsSH, SHsNH2, SHssNH, SHaaNH, SHsNH3p, SHssNH2p, SHsssNHp, SHtCH, SHdCH2, SHdsCH, SHaaCH, SHCHnX, SHCsats, SHCsatu, SHAVin, SHother, SHmisc, SsLi, SssBe, SssssBem, SsBH2, SssBH, SsssB, SssssBm, SsCH3, SdCH2, SssCH2, StCH, SdsCH, SaaCH, SsssCH, SddC, StsC, SdssC, SaasC, SaaaC, SssssC, SsNH3p, SsNH2, SssNH2p, SdNH, SssNH, SaaNH, StN, SsssNHp, SdsN, SaaN, SsssN, SddsN, SaasN, SssssNp, SsOH, SdO, SssO, SaaO, SaOm, SsOm, SsF, SsSiH3, SssSiH2, SsssSiH, SssssSi, SsPH2, SssPH, SsssP, SdsssP, SddsP, SsssssP, SsSH, SdS, SssS, SaaS, SdssS, SddssS, SssssssS, SSm, SsCl, SsGeH3, SssGeH2, SsssGeH, SssssGe, SsAsH2, SssAsH, SsssAs, SdsssAs, SddsAs, SsssssAs, SsSeH, SdSe, SssSe, SaaSe, SdssSe, SssssssSe, SddssSe, SsBr, SsSnH3, SssSnH2, SsssSnH, SssssSn, Ssl, SsPbH3, SssPbH2, SsssPbH, SssssPb, minHBd, minwHBd, minHBa, minwHBa, minHBint2, minHBint3, minHBint4, minHBint5, minHBint6, minHBint7, minHBint8, minHBint9, minHBint10, minHsOH, minHdNH, minHsSH, minHsNH2, minHssNH, minHaaNH, minHsNH3p, minHssNH2p, minHsssNHp, minHtCH, minHdCH2, minHdsCH, minHaaCH, minHCHnX, minHCsats, minHCsatu, minHAVin, minHother, minHmisc, minsLi, minssBe, minssssBem, minsBH2, minssBH, minsssB, minssssBm, minsCH3, mindCH2, minssCH2, mintCH, mindsCH, minaaCH, minsssCH, minddC, mintsC, mindssC, minaaC, minaaC, minssssC, minsNH3p, minsNH2, minssNH2p, mindNH, minssNH, minaaNH, mintN, minssssNHp, mindsN, minaaN, minsssN, minddsN, minaaN, minssssNp, minsOH, mindO, minssO, minaaO, minaOm, minsOm, minsF, minsSiH3, minssSiH2, minssssSiH, minsssssSi, minsPH2, minssPH, minsssP, mindsssP, minddsP, minsssssP, minsSH, mindS, minssS, minaaS, mindssS, minddssS, minssssssS, minSm, minsCl, minsGeH3, minssGeH2, minsssGeH, minssssGe, minsAsH2, minssAsH, minsssAs, mindsssAs, minddsAs, minsssssAs, minsSeH, mindSe, minssSe, minaaSe, mindssSe, minssssssSe, minddssSe, minsBr, minsSnH3, minssSnH2, minsssSnH, minssssSn, minsl, minsPbH3, minssPbH2,</p>		
--	---	--	--

	minssPbH, minssPb, maxHBd, maxwHBd, maxHBa, maxwHBa, maxHBint2, maxHBint3, maxHBint4, maxHBint5, maxHBint6, maxHBint7, maxHBint8, maxHBint9, maxHBint10, maxHsOH, maxHdNH, maxHsSH, maxHsNH2, maxHssNH, maxHaaNH, maxHsNH3p, maxHssNH2p, maxHsssNHp, maxHtCH, maxHdCH2, maxHdsCH, maxHaaCH, maxHCHnX, maxHCsats, maxHCsatu, maxHAvin, maxHother, maxHmisc, maxsLi, maxssBe, maxssBem, maxsBH2, maxssBH, maxssB, maxssBm, maxsCH3, maxdCH2, maxssCH2, maxtCH, maxdsCH, maxaaCH, maxssCH, maxddC, maxtsC, maxdssC, maxaaC, maxaaaC, maxssC, maxsNH3p, maxsNH2, maxssNH2p, maxdNH, maxssNH, maxaaNH, maxtN, maxssNHp, maxdsN, maxaaN, maxssN, maxdsN, maxaaN, maxssNp, maxsOH, maxdO, maxsO, maxaaO, maxaOm, maxsOm, maxsF, maxsSiH3, maxssSiH2, maxssSiH, maxssSi, maxsPH2, maxssPH, maxssP, maxdssP, maxdsP, maxssP, maxsSH, maxdS, maxssS, maxaaS, maxdssS, maxdssS, maxssS, maxSm, maxsCl, maxsGeH3, maxssGeH2, maxssGeH, maxssGe, maxsAsH2, maxssAsH, maxssAs, maxdssAs, maxdsAs, maxssAs, maxsSeH, maxdSe, maxssSe, maxaaSe, maxdssSe, maxssSe, maxdsSe, maxsBr, maxsSnH3, maxssSnH2, maxssSnH, maxssSn, maxsI, maxsPbH3, maxssPbH2, maxssPbH, maxssPb, sumI, meanI, hmax, gmax, hmin, gmin, LipoaffinityIndex, MAXDN, MAXDP, DELS, MAXDN2, MAXDP2, DELS2			
Extended topochemical atom	ETA_Alpha, ETA_AlphaP, ETA_dAlpha_A, ETA_dAlpha_B, ETA_Epsilon_1, ETA_Epsilon_2, ETA_Epsilon_3, ETA_Epsilon_4, ETA_Epsilon_5, ETA_dEpsilon_A, ETA_dEpsilon_B, ETA_dEpsilon_C, ETA_dEpsilon_D, ETA_Psi_1, ETA_dPsi_A, ETA_dPsi_B, ETA_Shape_P, ETA_Shape_Y, ETA_Shape_X, ETA_Beta, ETA_BetaP, ETA_Beta_s, ETA_BetaP_s, ETA_Beta_ns, ETA_BetaP_ns, ETA_dBeta, ETA_dBetaP, ETA_Beta_ns_d, ETA_BetaP_ns_d, ETA_Eta, ETA_EtaP, ETA_Eta_R, ETA_Eta_F, ETA_EtaP_F, ETA_Eta_L, ETA_EtaP_L, ETA_Eta_R_L, ETA_Eta_F_L, ETA_EtaP_F_L, ETA_Eta_B, ETA_EtaP_B, ETA_Eta_B_RC, ETA_EtaP_B_RC	FALSE	FALSE	TRUE
Hbond acceptor count	nHBAcc, nHBAcc2, nHBAcc3, nHBAcc_Lipinski	TRUE	FALSE	TRUE
Hbond donor count	nHBDon, nHBDon_Lipinski	TRUE	FALSE	TRUE
Hybridization ratio	HybRatio	TRUE	FALSE	TRUE
Kappa shape indices	Kier1, Kier2, Kier3	FALSE	TRUE	FALSE
Largest chain	nAtomLC	FALSE	TRUE	FALSE
Largest Pi system	nAtomP	FALSE	TRUE	FALSE

Longest aliphatic chain	nAtomLAC	TRUE	TRUE	FALSE
Mannhold LogP	MLogP	TRUE	TRUE	FALSE
McGowan volume	McGowan_Volume	FALSE	TRUE	FALSE
Molecular distance edge	MDEC-11, MDEC-12, MDEC-13, MDEC-14, MDEC-22, MDEC-23, MDEC-24, MDEC-33, MDEC-34, MDEC-44, MDEO-11, MDEO-12, MDEO-22, MDEN-11, MDEN-12, MDEN-13, MDEN-22, MDEN-23, MDEN-33	FALSE	TRUE	FALSE
Molecular linear free energy relation	MLFER_A, MLFER_BH, MLFER_BO, MLFER_S, MLFER_E, MLFER_L	FALSE	TRUE	FALSE
Petitjean number	PetitjeanNumber	FALSE	TRUE	FALSE
Ring count	nRing, n3Ring, n4Ring, n5Ring, n6Ring, n7Ring, n8Ring, n9Ring, n10Ring, n11Ring, n12Ring, nG12Ring, nFRing, nF4Ring, nF5Ring, nF6Ring, nF7Ring, nF8Ring, nF9Ring, nF10Ring, nF11Ring, nF12Ring, nFG12Ring, nHeteroRing, n3HeteroRing, n4HeteroRing, n5HeteroRing, n6HeteroRing, n7HeteroRing, n8HeteroRing, n9HeteroRing, n10HeteroRing, n11HeteroRing, n12HeteroRing, nG12HeteroRing, nFHeteroRing, nF4HeteroRing, nF5HeteroRing, nF6HeteroRing, nF7HeteroRing, nF8HeteroRing, nF9HeteroRing, nF10HeteroRing, nF11HeteroRing, nF12HeteroRing, nFG12HeteroRing, nTHeteroRing, nT4HeteroRing, nT5HeteroRing, nT6HeteroRing, nT7HeteroRing, nT8HeteroRing, nT9HeteroRing, nT10HeteroRing, nT11HeteroRing, nT12HeteroRing, nTG12HeteroRing	TRUE	TRUE	FALSE
Rotatable bonds count	nRotB, RotBFrac, nRotBt, RotBtFrac	TRUE	TRUE	FALSE
Topological	topoRadius, topoDiameter, topoShape	TRUE	TRUE	FALSE
Topological charge	GGI1, GGI2, GGI3, GGI4, GGI5, GGI6, GGI7, GGI8, GGI9, GGI10, JGI1, JGI2, JGI3, JGI4, JGI5, JGI6, JGI7, JGI8, JGI9, JGI10, JGT	FALSE	TRUE	FALSE
Topological distance matrix	SpMax_D, SpDiam_D, SpAD_D, SpMAD_D, EE_D, VE1_D, VE2_D, VE3_D, VR3_D	FALSE	TRUE	FALSE
Topological polar surface area	TopoPSA	TRUE	TRUE	FALSE
Van der Waals volume	VABC	TRUE	TRUE	FALSE
Weight	MW, AMW	TRUE	TRUE	FALSE
Wiener numbers	WPATH, WPOL	FALSE	FALSE	TRUE
XLogP	XLogP	FALSE	FALSE	TRUE

Tabel 4: fingerprints die van belang zijn voor oxidatieprocessen (rood), membraanfiltratie (geel) of beide (groen)

Fingerprint	Van belang voor oxidatie	Van belang voor membraanfiltratie	Van belang voor zowel oxidatie als membraanfiltratie
Primary carbon	X		
Secondary carbon	X		
Tertiary carbon	X		
Quaternary carbon	X		
Alkene	X		
Alkyne	X		
Allene	X		
Alkylchloride	X		
Alkylfluoride	X		
Alkylbromide	X		
Alkyljodide	X		
Alcohol	X		
Primary alcohol	X		
Secondary alcohol	X		
Tertiary alcohol	X		
Dialkylether	X		
Dialkylthioether	X		
Alkylarylether	X		
Diarylether	X		
Alkylarylthioether	X		
Diarylthioether	X		
Oxonium	x		
Amine			X
Primary aliph amine			X
Secondary aliph amine			X
Tertiary aliph amine			X
Quaternary aliph ammonium			X
Primary arom amine			X
Secondary arom amine			X
Tertiary arom amine			X
Quaternary arom ammonium			X
Secondary mixed amine			X
Tertiary mixed amine			X
Quaternary mixed ammonium			x
Ammonium	X		
Alkylthiol	X		
Dialkylthioether	X		
Alkylarylthioether	X		
Disulfide	X		
1,2-Aminoalcohol	X		

1,2-Diol	X		
1,1-Diol	X		
Hydroperoxide	X		
Peroxo	X		
Aldehyde	X		
Ketone	X		
Thioaldehyde	X		
Thioketone	X		
Imine	X		
Immonium	X		
Oxime	X		
Oximether	X		
Acetal	X		
Hemiacetal	x		
Aminal	X		
Hemiaminal	X		
Thioacetal	X		
Thiohemiacetal	X		
Halogen acetal like	X		
Acetal like	X		
Halogenmethylen ester and similar	X		
NOS methylen ester and similar	X		
Hetero methylen ester and similar	X		
Cyanhydrine	X		
Chloroalkene	X		
Fluoroalkene	X		
Bromoalkene	X		
Iodoalkene	X		
Enol	X		
Endiol	X		
Enolether	X		
Enolester	X		
Enamine	X		
Thioenol	X		
Thioenolether	X		
Acylchloride	X		
Acylfluoride	X		
Acylbromide	X		
Acyliodide	X		
Acylhalide	X		
Carboxylic acid			x
Carboxylic ester	X		

Lactone	X		
Carboxylic anhydride	X		
Carboxylic acid derivative	X		
Carbothioic acid	X		
Carbothioic S ester	X		
Carbothioic S lactone	X		
Carbothioic O ester	x		
Carbothioic O lactone	X		
Carbothioic halide	X		
Carbodithioic acid	X		
Carbodithioic ester	X		
Carbodithiolactone	X		
Amide	X		
Primary amide	X		
Secondary amide	X		
Tertiary amide	X		
Lactam	X		
Alkyl imide	x		
N hetero imide			X
Imide acidic			X
Thioamide	x		
Thiolactam	X		
Oximester	x		
Amidine			X
Hydroxamic acid			X
Hydroxamic acid ester	x		
Imidoacid			X
Imidoacid cyclic			X
Imidoester	X		
Imidolactone	X		
Amidine			X
Imidolactam			X
Imidoylhalide			X
Imidoylhalide cyclic			x
Amidrazone	x		
Alpha aminoacid			X
Alpha hydroxyacid			X
Peptide middle	X		
Peptide C term	X		
Peptide N term	X		
Carboxylic orthoester	X		
Ketene	X		
Ketenacetal	X		
Nitrile	X		

Isonitrile	X		
Vinylogous carbonyl or carboxyl derivative	x		
Vinylogous acid			x
Vinylogous ester	x		
Vinylogous amide			x
Vinylogous halide	X		
Carbonic acid dieester	X		
Carbonic acid esterhalide	X		
Carbonic acid monoester	x		
Carbonic acid derivatives			x
Thiocarbonic acid dieester	X		
Thiocarbonic acid esterhalide	X		
Thiocarbonic acid monoester	X		
Urea	X		
Thiourea	X		
Isourea	X		
Isothiourea	x		
Guanidine			X
Carbaminic acid			X
Urethan			X
Biuret			X
Semicarbazide			X
Carbazide			X
Semicarbazone			X
Carbazone			X
Thiosemicarbazide			X
Thiocarbazide			X
Thiosemicarbazone			X
Thiocarbazone			x
Isocyanate	X		
Cyanate	X		
Isothiocyanate	X		
Thiocyanate	X		
Carbodiimide	x		
Orthocarbonic derivatives			X
Phenol			X
1,2-Diphenol			X
Arylchloride			X
Arylfluoride			X
Arylbromide			X
Aryliodide			X
Arylthiol			X
Iminoarene			X

Oxoarene			X
Thioarene			X
Hetero N basic H			X
Hetero N basic no H			X
Hetero N nonbasic			x
Hetero O	X		
Hetero S	X		
Heteroaromatic	X		
Nitrite	X		
Thionitrite	X		
Nitrate	X		
Nitro	X		
Nitroso	X		
Azide	X		
Acylazide	X		
Diazo	X		
Diazonium	X		
Nitrosamine	X		
Nitrosamide	X		
N-Oxide	X		
Hydrazine	X		
Hydrazone			X
Hydroxylamine			X
Sulfon	X		
Sulfoxide	X		
Sulfonium	X		
Sulfuric acid			x
Sulfuric monoester	x		
Sulfuric diester	x		
Sulfuric monoamide			X
Sulfuric diamide			X
Sulfuric esteramide			X
Sulfonic acid			X
Sulfonamide			x
Sulfonic ester	X		
Sulfonic halide	X		
Sulfonic derivative	x		
Sulfinic acid			X
Sulfinic amide			X
Sulfinic ester	X		
Sulfinic halide	X		
Sulfinic derivative	x		
Sulfenic acid			x
Sulfenic amide			x

Sulfenic ester	X		
Sulfenic halide	X		
Sulfenic derivative	X		
Phosphine	X		
Phosphine oxide	X		
Phosphonium	X		
Phosphorylen	X		
Phosphonic acid			x
Phosphonic monoester	X		
Phosphonic diester	X		
Phosphonic monoamide			X
Phosphonic diamide			X
Phosphonic esteramide			X
Phosphonic acid derivative			X
Phosphoric acid			X
Phosphoric monoester	X		
Phosphoric diester	X		
Phosphoric triester	x		
Phosphoric monoamide			X
Phosphoric diamide			X
Phosphoric triamide			X
Phosphoric monoestermonoamide			X
Phosphoric diestermonoamide			X
Phosphoric monoesterdiamide			X
Phosphoric acid derivative			X
Phosphinic acid			x
Phosphinic ester	x		
Phosphinous amide			x
Phosphinous derivatives	X		
Trialkylborane	X		
Boric acid derivatives	X		
Boronic acid derivative	x		
Borohydride		x	
Aromatic			x
Heterocyclic	X		
Epoxide	X		
NH aziridine	X		
Spiro	x		
Annelated rings		X	
Bridged rings		x	

Sugar pattern 1	X		
Sugar pattern 2	X		
Sugar pattern combi	X		
Sugar pattern 2 reducing	X		
Sugar pattern 2 alpha	X		
Sugar pattern 2 beta	X		
Conjugated double bond	X		
Conjugated tripple bond	X		
Cis double bond	X		
Trans double bond	X		
Mixed anhydrides	x		
Trifluoromethyl			x
Charged		X	
Anion		X	
Kation		X	
Salt		X	
Michael acceptor	x		
Dicarbodiazene			x
CH-acidic		X	
CH-acidic strong		X	