

Bedrijfstakonderzoek
BTO 2024.051 | April 2024

Non-target screening: kwantitatief en op tijd

Bedrijfstakonderzoek

KWR

Bridging Science to Practice

Rapport

Non-target screening: kwantitatief en op tijd

BTO 2024.051 | April 2024

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

Opdrachtnummer

402045/151

Projectmanager

Dr. Patrick S. Bäuerlein

Opdrachtgever

BTO - Thematisch onderzoek - Chemische veiligheid

Auteur(s)

Nienke Meekel MSc, Dennis Vughs MSc, Patrick Ippel, Dr. Andrea Brunner

Kwaliteitsborger(s)

Dr. Thomas ter Laak, Dr. Frederic Béen

Verzonden naar

Dit rapport is verspreid onder BTO-participanten.

Een jaar na publicatie is het openbaar.

Keywords

Non-target screening, chemische veiligheid, chemische waterkwaliteit, massaspectrometrie, monitoring, LC-HRMS, organische microverontreinigingen

Jaar van publicatie
2024

Meer informatie
Nienke Meekel MSc
T 030-6069622
E nienke.meekel@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

April 2024 ©

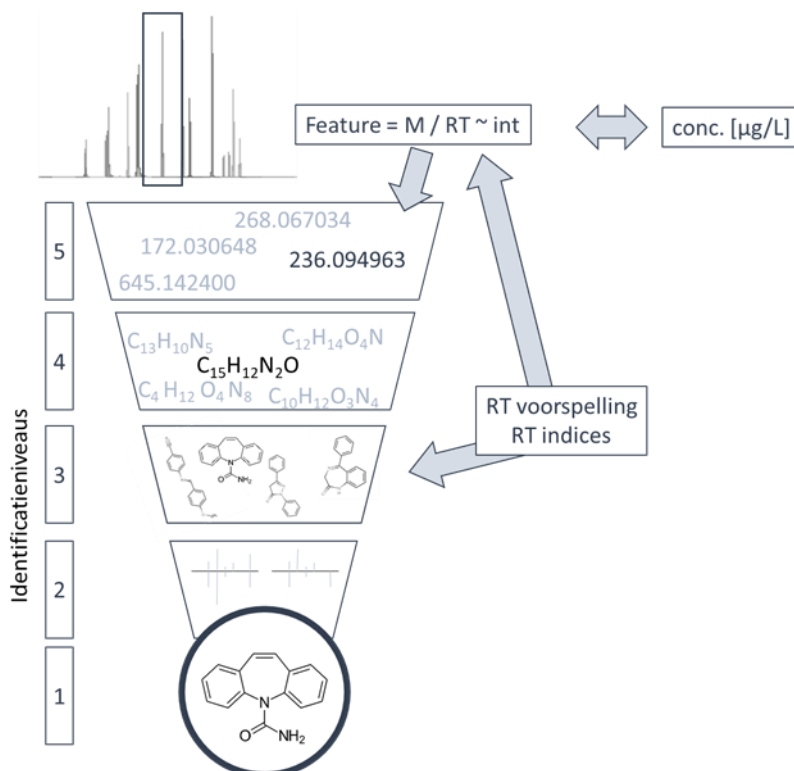
Alle rechten voorbehouden aan KWR. Niets uit deze uitgave mag - zonder voorafgaande schriftelijke toestemming van KWR - worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen,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

Managementsamenvatting

Retentietijdvoorspelling is nu nog niet volwassen, maar semi-kwantificatie is al veelbelovend voor vergemakkelijken van interpretatie van non-target screening naar onbekende microverontreinigingen in water

Nienke Meekel, Dennis Vughs, Patrick Ippel en Andrea Brunner

De ontwikkelingen van retentietijdvoorspellende modellen en kwantificatie zonder referentiestandaarden gaan hard. In dit onderzoek zijn diverse beschikbare modellen getest. De retentietijdvoorspellende modellen bleken nog niet voldoende nauwkeurig te zijn om ze toe te passen bij non-target screening. De modellen voor semikwantificatie presteren beter. Het advies is dan ook om het beste semi-kwantificatie model Quantem toe te passen waar dat relevant is voor de onderzoeksvraag en in de komende jaren de (snelle) ontwikkeling rond beide soorten modellen goed te volgen en in te zetten zo gauw dat zinvol is.



Schematisch overzicht van identificatie en kwantificering op basis van non-target screening

Belang: uitbreiding van NTS workflows voor identificatie van waterverontreiniging

Non-target screening (NTS) wordt veelvuldig toegepast voor de detectie en identificatie van onbekende waterverontreinigingen. In dit project is onderzocht of de NTS-workflow kan worden uitgebreid met retentietijdvoorspellende modellen en kwantificatie zonder referentiestandaarden. Door

met retentietijdvoorspellende modellen de retentietijden van kandidaatstoffen (suspects) te voorspellen, kan het aantal suspects per te identificeren stof worden verminderd. Semi-kwantitatieve modellen kunnen worden ingezet om de concentratie van een onbekende stof te voorspellen. Voor antropogene stoffen geldt in Nederland een drempelwaarde van 1 µg/L in

oppervlaktewater bestemd voor de bereiding van drinkwater. Wanneer aan een onbekende stof een concentratie kan worden toegekend zonder referentiestandaarden te gebruiken, kan dat een belangrijke bijdrage leveren aan de prioritering en risicobeoordeling van die stof.

Aanpak: Vergelijk retentietijdvoorspellende modellen en semi-kwantitatieve modellen

Zeven retentietijdvoorspellende modellen van uiteenlopende complexiteit zijn met elkaar vergeleken. De drie best presterende retentietijdvoorspellende modellen zijn toegepast op een NTS dataset, waarna vervolgens hun prestaties vergeleken zijn. Daarnaast zijn via deelname aan het NORMAN ringonderzoek voor (semi-)kwantitatieve LC-HRMS non-target screening, verschillende semi-kwantificatie modellen getest. Het commerciële en meest gebruiksvriendelijke model Quantem is uitgebreider getest en vergeleken met de huidige methode om de concentraties te schatten als interne standaard equivalenten.

Resultaten: Retentietijdvoorspelling nog niet volwassen, semi-kwantificatie veelbelovend

De meest veelbelovende retentietijdvoorspellende modellen zijn QSRR Automator en Retip. De prestaties van deze modellen zijn echter nog niet zodanig dat ze een toegevoegde waarde hebben in de reguliere NTS workflows. Momenteel lijken deze modellen nog te vaak de daadwerkelijke stof uit de kandidatenlijst te filteren. De verwachting is dat hun toegevoegde waarde in de toekomst zal verbeteren, omdat de ontwikkeling van deze modellen voortgaat en er steeds meer en betere trainingdata beschikbaar komt. Het meest veelbelovende semi-kwantitatieve model is Quantem. Quantem lijkt betere resultaten op te leveren dan de referentiestandaard equivalenten.

Toepassing: Aan de slag met semi-kwantificatie

Momenteel lijken de retentietijdvoorspellende modellen nog te vaak de daadwerkelijke stof uit de kandidatenlijst te filteren. Het advies is om over een paar jaar de stand van zaken te heroverwegen vanwege de snelgroeiende ontwikkelingen. De semikwantificatie is verder ontwikkeld, maar het gebruik van Quantem brengt kosten met zich mee. Daarom is het advies om Quantem alleen toe te passen bij projecten of vraagstukken waar de semi-kwantificatie echt van toegevoegde waarde is. De semi-kwantitatieve modellen voor non-target en suspect screening zijn volop in ontwikkeling en de verwachting is dat deze in de komende jaren sterk zullen verbeteren. Ten slotte kunnen de verder ontwikkelde en nieuw ontwikkelde semi-kwantitatieve tools nader onderzocht worden zodra de resultaten van het NORMAN ringonderzoek voor (semi-)kwantitatieve LC-HRMS non-target screening bekend zijn en de aanbevelingen daaruit ook in ogenschouw genomen kunnen worden.

Rapport

Dit onderzoek is beschreven in het rapport *Non-target screening: kwantitatief en op tijd* (BTO-2024.051).

Inhoud

Rapport	2
<i>Managementsamenvatting</i>	3
Inhoud	5
Lijst van afkortingen	7
1 Introductie	8
1.1 Retentietijd indices	9
1.2 Ionisatie-efficiëntie	9
2 Retentietijdvoorspelling en retentietijd indices	10
2.1 Lineaire regressie: logP en logD	11
2.2 QSAR modellen: OPERA-RT, QSRR Automator, QSRR Athene en Retip	11
2.2.1 OPERA-RT en combinatiemodel	12
2.2.2 QSRR Automator	12
2.2.3 QSRR Athene	12
2.2.4 Retip	13
2.3 Vergelijking modelprestaties	13
2.4 LC-HRMS analyse	14
2.4.1 Chemicaliën	14
2.4.2 LC-HRMS	14
2.4.3 Bepaling RT en opname MS2 spectra	14
2.5 Resultaten retentietijdvoorspelling	14
2.5.1 Verdeling van de data	14
2.5.2 LogP, logD en OPERA-RT modellen (lineaire- en meervoudige-lineaire regressie)	15
2.5.3 QSRR modellen	18
2.6 Modelprestaties in de praktijk	23
2.7 Discussie van de modelprestaties	26
2.8 Toepassing op NTS dataset	27
2.9 Gebruiksgemak en koppeling met patRoom	29
2.10 Conclusie retentietijdvoorspellende modellen	29
3 Semikwantitatieve NTS	30
3.1 Methoden om ionisatie-efficiëntie te voorspellen	31
3.1.1 Vergelijkbare stoffen	31
3.1.2 Nabij eluerende stof	31
3.1.3 Voorspellen van ionisatie-efficiëntie	31
3.2 NORMAN ringonderzoek voor (semi-)kwantitatieve LC-HRMS non-target screening	32

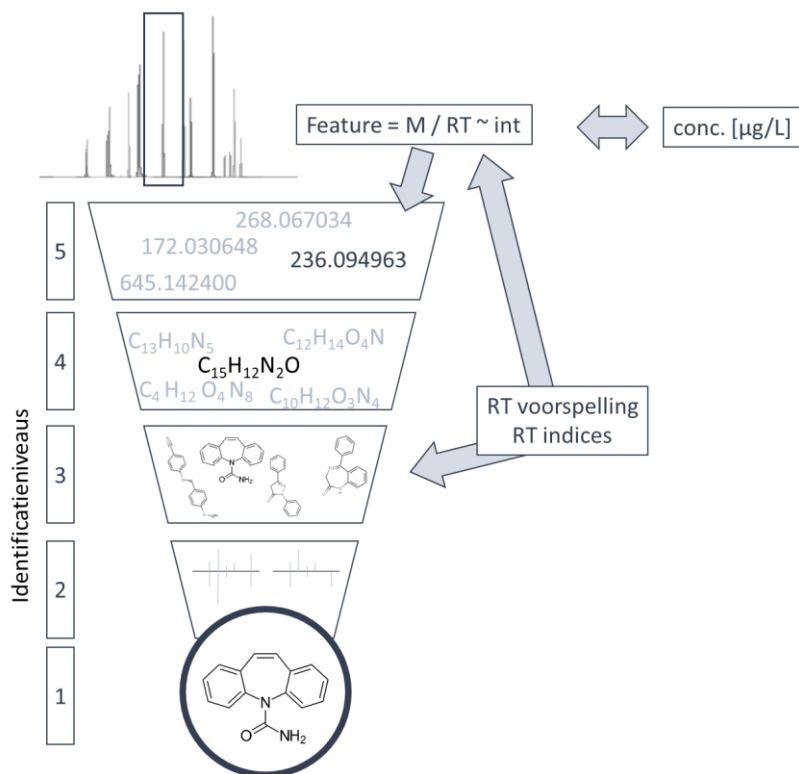
3.3	Test Quantem	33
3.3.1	Methode	33
3.3.2	Resultaten	33
3.3.3	Discussie	38
3.4	Conclusie semi-kwantitatieve modellen	39
4	Referenties	40
I	Overzicht stoffen	42
II	Resultaten Quantem	45

Lijst van afkortingen

BRNN	Bayesian-regularized neural network
ESI	electrospray ionisatie
DW	drinkwater
EFF	rioolwaterzuiveringinstallatie effluent
FWHM	Full Width at Half Maximum
GC	gaschromatografie
HRMS	hoge resolutie massaspectrometrie
IE	ionisatie-efficiëntie
LightGBM	light gradient boosting machine
LC	vloeistofchromatografie
LogP	octanol-water partitie coëfficiënt
LogD	octanol-water partitie coëfficiënt bij een vaste pH
MAE	mean absolute error
MAPE	mean absolute percentage error
MeOH	methanol
MS	massaspectrometrie
MS2	fragmentatiespectrum
MQ	milli-Q water
m/z	massa-ladings ratio
NTS	non-target screening
OMV's	organische microverontreinigingen
OPERA	OPEn structure-activity/property Relationship App
OW	oppervlaktewater
PC	Principal Component
PCA	Principal Component Analysis (Hoofdcomponentenanalyse)
pKa	zuurconstante
pKb	baseconstante
QSAR	Quantitative Structure-Activity Relationship
QSRR	Quantitative Structure-Retention Relationship
Q ² _{Loo}	leave-one out kruisvalidatie correlatiecoëfficiënt
RF	Random Forest
RMSE	root mean square error
RT	retentietijd
RTI	retentietijdindex
SMILES	simplified molecular-input line-entry system

1 Introductie

De risicobeoordeling van een onbekende stof vereist betrouwbare identificatie van het molecuul, gevolgd door de accurate bepaling van de concentratie in het watermonster. In de laatste jaren is de monitoring van de chemische waterkwaliteit ontwikkeld van doelstofanalyses (gericht op het identificeren en kwantificeren van geselecteerde stoffen) naar suspect en non-target screening (NTS, gericht op het identificeren van de 'bekende onbekende' en 'onbekende onbekenden'). NTS op basis van vloeistofchromatografie gekoppeld aan elektrospay ionisatie hoge resolutie massaspectrometrie (LC ESI-HRMS) kan een breed spectrum aan ioniseerbare stoffen detecteren die scheidbaar zijn met de geselecteerde chromatografie. Uiteindelijk geeft dit dus een completer beeld van de aanwezigheid van organische microverontreinigingen (OMV's) in water dan met doelstofanalyses.¹ De identificatie van stoffen uit NTS data is uitdagend vanwege de grote hoeveelheid potentiële kandidaatstoffen, de benodigde tijd om tot een identificatie te komen en de onzekerheden bij een tentatieve identificatie. Daar komt bij dat NTS niet kwantitatief is: intensiteiten en oppervlakten van pieken kunnen niet direct vertaald worden naar stofconcentraties door de verschillen in ionisatie-efficiëntie (IE). Anders dan bij doelstofanalyses kan hier in NTS niet voor gecorrigeerd worden met beschikbare zuivere standaarden of bekende relaties. Momenteel is de kwantificering van OMV's afhankelijk van het gebruik van referentiestandaarden. In dit project worden deze twee uitdagingen (hoeveelheid kandidaatstoffen en kwantificering) geadresseerd om zo het gat tussen doelstofanalyses en non-target screening kleiner te maken.



Figuur 1 - Schematisch overzicht van identificatie en kwantificering op basis van non-target screening.

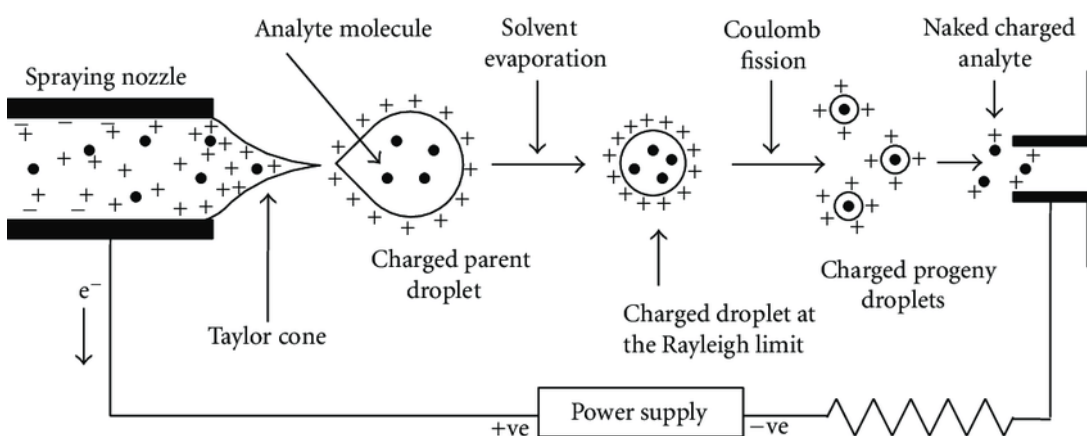
1.1 Retentietijd indices

Retentietijden kunnen worden gebruikt voor de identificatie van een stof. Echter, de retentietijd is afhankelijk van de chromatografische methode waardoor niet alle retentietijden met elkaar vergeleken kunnen worden. Een oplossing hiervoor is de retentietijdindex (RTI), veelvuldig gebruikt in GC-MS, waarmee de retentietijd gecorrigeerd/genormaliseerd wordt op basis van de retentietijden van één of meer interne standaarden.² Met behulp van de retentietijdindex kan per laboratorium de verwachte retentietijd van de stof in de eigen chromatografische methode berekend worden. De retentietijdindex kan tevens worden gebruikt om resultaten van verschillende laboratoria met elkaar te vergelijken. Er zijn verschillende modellen ontwikkeld om de retentietijd index en/of de retentietijd van een onbekende stof te voorspellen op basis van de eigenschappen van de stof. Aan de hand van deze voorspelde retentietijd kunnen de kandidaatstoffen worden gefilterd, zodat de kandidaten die vergelijkbaar zijn met de te onderzoeken stof overblijven.

1.2 Ionisatie-efficiëntie

Het is van belang om de intensiteit van een NTS feature om te zetten in een concentratie, de mate waarin een stof aanwezig is in een matrix is een vereiste voor een volledige risicobeoordeling. Een voorbeeldtoepassing is de algemene streefwaarde met alarmfunctie (de zogeheten signaleringsparameter) voor oppervlaktewater bestemd voor de bereiding van drinkwater welke is vastgesteld op $1 \mu\text{g/L}$,³ en de signaleringswaarde van $0,1 \mu\text{g/L}$ ⁴ voor "andere" organische antropogene stoffen in bronnen voor drinkwater in Nederland. Ionisatie-efficiëntie kan worden gebruikt om op basis van de oppervlakte of intensiteit van een piek (d.w.z. een feature: combinatie van retentietijd, m/z waarde en intensiteit), de concentratie van de bijbehorende stof te voorspellen.

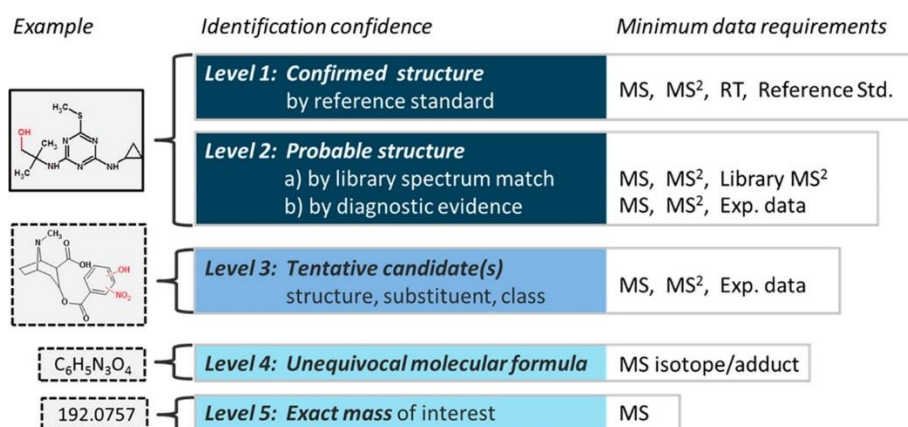
Electrospray ionisatie (ESI) is de meest gangbare ionisatietechniek voor LC-MS. ESI is een zachte ionisatietechniek waarbij gebruik gemaakt wordt van een sterk elektrisch veld tussen een capillair en een counter-elektrode, zie Figuur 2. Hierdoor ontstaat een ophoping van lading aan het uiteinde van het capillair, waar de stof uit komt. Vervolgens breekt dit vloeistofoppervlakte aan het einde van het capillair uiteen in geladen druppeltjes, waaruit het oplosmiddel verdampt en waar vervolgens de overgebleven geladen deeltjes verder de massaspectrometer in gaan. ESI kan worden uitgevoerd in positieve en negatieve modus, hierbij ontstaan respectievelijk positief $[M+H]^+$ en negatief geladen ionen $[M-H]^-$. De mate waarin moleculen in de vloeistoffase omgezet worden naar ionen in de gasfase en uiteindelijk gedetecteerd worden in de detector, wordt aangeduid met de term 'ionisatie-efficiëntie' (IE).⁵ Als het mogelijk is om de IE van een bepaalde stof met voldoende nauwkeurigheid te voorspellen, dan kan de piekoppervlakte vertaald worden naar een concentratie (schatting).



Figuur 2 - Schematische weergave van het electrospray ionisatie proces, figuur afkomstig uit Banerjee and Mazumdar⁶.

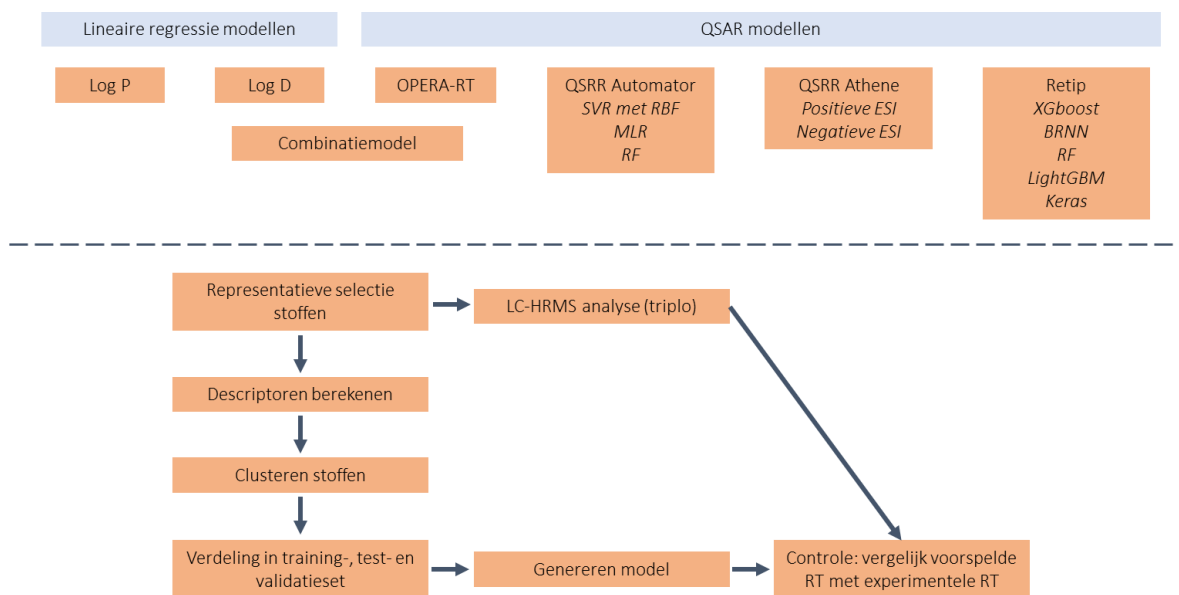
2 Retentietijdvoorspelling en retentietijd indices

Om een onbekende stof te identificeren worden er verschillende stappen doorlopen. Uit een LC-HRMS analyse worden 'features' verkregen: signalen die bestaan uit een retentietijd, intensiteit en accurate massa. Schymanski, Jeon, Gulde, Fenner, Ruff, Singer and Hollender ⁷ ontwikkelde een systeem bestaande uit verschillende identificatieniveaus (Figuur 3) waarbij op basis van verschillende eigenschappen (accurate massa, isotopenpatroon, fragmentatiespectrum, retentietijd, database spectrum en een referentiestandaard) kandidaten worden gezocht, de beschikbare gegevens bepalen welk identificatieniveau erbij hoort. Vanaf niveau 3 (Figuur 3) worden features uitgedrukt in mogelijke chemische structuren, waarbij meerdere structuren (kandidaten) passen bij het beschikbare bewijs. Hoe meer bewijs beschikbaar is, hoe nauwer de lijst van kandidaten wordt. Een kandidaat, kandidaatstructuur of stofkandidaat is dus een chemische structuur die overeen zou kunnen komen met de betreffende feature. De retentietijd óf voorspelde retentietijd kan gebruikt worden om het aantal kandidaatstructuren voor een feature te reduceren. Kandidaten met een veel kortere of langere (voorspelde) retentietijd dan de feature in kwestie zijn immers geen reële optie. In Figuur 3 is de retentietijd pas weergegeven bij niveau 1 'bevestigde structuur', in de praktijk speelt de retentietijd op lagere niveaus (niveau 2 en 3) ook een rol.



Figuur 3 - Betrouwbaarheidsniveaus in HRMS analyse van Schymanski, Jeon, Gulde, Fenner, Ruff, Singer and Hollender ⁷

Er zijn verschillende modellen beschikbaar om de retentietijd van stofkandidaten te voorspellen op basis van verschillende stoffeigenschappen. Deze modellen lopen uiteen van relatief simpel met behulp van lineaire regressie (bijv. op basis van logP, de octanol-water partitie coëfficiënt), tot vrij complex met behulp van bijvoorbeeld kunstmatige neurale netwerken en support vector machines (bijv. op basis van een veelvoud aan eigenschappen). In dit project zijn de modellen Retip⁸, QSRR automator⁹, QSRR Athene¹⁰, logP, logD en OPERA-RT¹¹ met elkaar vergeleken, zie Figuur 4. De modellen zijn op verschillende manieren opgebouwd en verschillen daarin ook in complexiteit. De stappen die voorafgaan aan het genereren van een retentietijdvoorspellend model zijn weergegeven in Figuur 4.



Figuur 4 - Boven: overzicht van de verschillende geteste modellen. Onder: Verschillende stappen in de ontwikkeling van een retentietijdvoorspellend model.

2.1 Lineaire regressie: logP en logD

Twee voorbeelden van lineaire regressiemodellen zijn het logP- en het logD-model. In beide modellen wordt de correlatie tussen de variabele (logP of logD) en de experimentele retentietijd gebruikt om de verwachte retentietijd van de onbekende stof te voorspellen. De logP en logD geven de octanol-water partitie coëfficiënt weer, waarbij in het geval van de logD een vaste pH geldt (hier wordt rekening gehouden met de betreffende ioniserende vormen van het molecuul, in tegenstelling tot de logP waarbij dit niet gebeurt). Deze parameters kunnen experimenteel bepaald worden of theoretisch met behulp van software zoals EPI Suite^{TM 12} of OPERA¹³. Verschillende databases zoals ChemSpider (<http://www.chemspider.com/>) en U.S. EPA CompTox Chemicals Dashboard (<https://comptox.epa.gov/dashboard>) bevatten informatie over deze parameters. Voor het logP model zijn verschillende logP waarden verzameld; voorspeld door EPI SuiteTM, OPERA en ACD/Labs. Deze logP waarden zijn gemiddeld tot één logP waarde per stof. Voor het logD model is hetzelfde gedaan op basis van twee verschillende, door OPERA voorspelde, logD waarden (bij pH 5,5 en pH 7,4).¹ Vervolgens zijn de logP en logD van elke stof uitgezet tegen de experimentele retentietijden en zijn hier twee modellen van gemaakt met behulp van lineaire regressie.

2.2 QSAR modellen: OPERA-RT, QSRR Automator, QSRR Athene en Retip

De andere vier onderzochte modellen (OPERA-RT, QSRR Automator, QSRR Athene en Retip) zijn QSAR (Quantitative Structure-Activity Relationship) modellen. Dit type model wordt veelal gebruikt om onbekende eigenschappen (zoals toxiciteit) van een stof te voorspellen op basis van verschillende descriptoren die de structuur van de stof omschrijven. Voorbeelden van deze descriptoren zijn de logP waarde, het aantal zuurgroepen, het aantal koolstofatomen of het aantal draaiende bindingen. QSAR modellen kunnen zo ook gebruikt worden om de retentietijd te voorspellen, in dit geval worden ze QSRR (Quantitative Structure Retention Relationship) modellen genoemd.

¹ De gebruikte logD waarden komen niet overeen met de pH van het eluents. Idealiter is dit wel het geval, om zo een optimale vergelijking te kunnen maken tussen de voorspelde en de experimentele retentietijden, welke onderhevig zijn aan pH invloeden. De logD informatie bij pH 5,5 en pH 7,4 is vrij beschikbaar, de gegevens van een lagere pH kunnen berekend worden met behulp van commerciële software. Gezien de benodigde tijd en middelen, en de focus van dit onderzoek (welke niet op de logD lag, maar op de QSAR modellen) is ervoor gekozen om de vrij beschikbare data te gebruiken.

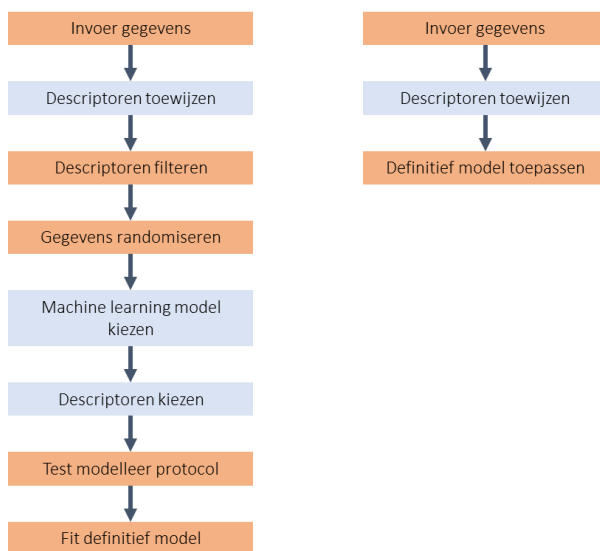
2.2.1 OPERA-RT en combinatiemodel

OPERA is een omgeving van QSAR modellen die gebruikt kan worden om verschillende eigenschappen te voorspellen.¹³ Het OPERA-RT model¹¹ (<https://github.com/kmansouri/OPERA>) is vergelijkbaar, maar dan gericht op het voorspellen van de retentietijd van een stof aan de hand van een veelvoud van moleculaire descriptoren welke berekend worden met de software PaDEL¹⁴. Na berekening worden de vrijwel constante descriptoren of waarvan de waarden ontbreken, verwijderd.¹¹ In het bijzonder, worden de best voorspellende descriptoren bepaald met behulp van genetische algoritmen (dit zijn optimalisatie-algoritmen die de meest geschikte combinatie van descriptoren zoeken om een bepaald probleem op te lossen, in dit geval retentietijdvoorspelling). Uiteindelijk ontstaat er zo een model waarin de nuttigste descriptoren worden meegenomen om zo de retentietijd van een stof te kunnen bepalen.

Het OPERA-RT model is momenteel gebaseerd op één chromatografisch systeem. Dit vereist dus een formule die de door OPERA-RT voorspelde retentietijd omzet naar de juiste retentietijd die wordt verwacht in het gebruikte chromatografische systeem. Met behulp van multiple lineaire regressie is het OPERA-RT model gecombineerd met het logP- en het logD model om zo de combinatie van deze drie variabelen te testen; het combinatiemodel.

2.2.2 QSRR Automator

QSRR Automator⁹ is een tool voor de ontwikkeling van een retentietijdvoorspellend model voor lipidomics en metabolomics. Dit kan op drie manieren: met behulp van support vector machines voor regressie (SVR) met een radiale basis functie kernel (RBF), multiple lineaire regressie (MLR) of random forest (RF). Om een model samen te stellen moet er een trainingset bestaande uit SMILES geüpload worden (Figuur 5), vervolgens kunnen er descriptoren ingevoerd worden door de gebruiker of berekend worden door de software zelf. Deze descriptoren worden gefilterd (duplicaten, ontbrekende waarden of met een correlatie > 0,9 t.o.v. andere descriptoren) en de data wordt gerandomiseerd. De gebruiker kan een machine learning model kiezen, of de software laten onderzoeken welk model de beste resultaten geeft. Ten slotte wordt het model gegenereerd en cross-validatie uitgevoerd. De QSRR Automator software is beschikbaar via [deze link](#).



Figuur 5 - QSRR Automator workflow⁹, links wordt de training van een model weergegeven en rechts de stappen in de retentietijdvoorspelling met behulp van het getrainde model. De blauwe stappen geven aan waar de gebruiker zelf instellingen kan invoeren.

2.2.3 QSRR Athene

Aalizadeh, Thomaidis, Bletsou and Gago-Ferrero¹⁰ hebben twee verschillende QSRR modellen ontwikkeld voor de voorspelling van de retentietijd in twee reversed-phase LC methodes; voor positieve- en negatieve elektrospay ionisatie. In beide LC methodes werd dezelfde C18 kolom gebruikt, de mobiele fase van de positieve ionisatie

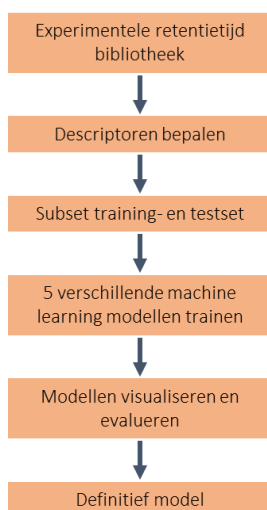
methode bestond uit (A) H₂O:MeOH (90:10) met 5 mM ammoniumformiaat en 0,01% mierenzuur en (B) MeOH 5 mM ammoniumformiaat en 0,01% mierenzuur (pH 3,6). De mobiele fase van de negatieve ionisatie methode bestond uit (A) H₂O:MeOH (90:10) met 5 mM ammoniumacetaat en (B) MeOH 5 mM ammoniumacetaat (pH 6,2). Om gebruik te maken van de modellen moeten er 36 referentiestandaarden (18 voor elke ionisatiemethode) gemeten worden en kan met behulp van de berekende retentietijdindex (vergelijking 1) de retentietijden voorspeld worden voor de kandidaatstoffen. De online software is beschikbaar via [deze link](#).

$$RTI = \frac{(RT_x - RT_{min})}{(RT_{max} - RT_{min})} \cdot 1000 \quad (1)$$

Beide modellen zijn gemaakt met behulp van support vector machine (SVM), dit algoritme valt onder supervised machine learning.

2.2.4 Retip

Een andere tool die gebruikt kan worden voor de ontwikkeling van een QSRR model is 'Retip'⁸ (<https://www.retip.app>), welke is uitgevoerd als een R-pakket. Retip maakt gebruik van vijf verschillende regressiemodellen: XGboost, Bayesian-regularized neural network (BRNN), Random Forest (RF), light gradient boosting machine (LightGBM) en Keras.⁸ In Figuur 6 is de standaard workflow van de software weergegeven waarbij de gebruiker in de eerste stap een bibliotheek met SMILES en retentietijden uploadt.



Figuur 6 - Stappenplan Retip⁸.

2.3 Vergelijking modelprestaties

Om de prestaties van de verschillende modellen met elkaar te vergelijken zijn een aantal parameters bepaald. De gemiddelde absolute fout (MAE, mean absolute error) van de experimentele versus de voorspelde retentietijd (2), de (MAPE, mean absolute percentage error) (3), de wortel van de gemiddelde gekwadraterde fout (RMSE, root mean square error) (4) en de leave-one out kruisvalidatie correlatiecoëfficiënt (Q²_{LOO}) (5). In deze vergelijkingen staat n voor het aantal gemeten stoffen en i geeft de individuele stof aan, van 1 tot n .

$$MAE = \frac{\sum_{i=1}^n \text{abs}(Voorspelde RT_i - Experimentele RT_i)}{n} \quad (2)$$

$$MAPE = \frac{\sum_{i=1}^n \text{abs}\left(\frac{Experimentele RT_i - Voorspelde RT_i}{Experimentele RT_i}\right) \cdot 100\%}{n} \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\text{Experimentele } RT_i - \text{Voorspelde } RT_i)^2}{n}} \quad (4)$$

$$Q_{Loo}^2 = r_{CV}^2 = 1 - \frac{\sum_i^l (\text{Experimentele } RT_i - \text{Voorspelde } RT_i)^2}{\sum_i^l (\text{Experimentele } RT_i - \text{Gemiddelde } RT_i)^2} \quad (5)$$

2.4 LC-HRMS analyse

2.4.1 Chemicaliën

Van 316 verschillende stoffen werd de retentietijd bepaald, de standaarden zijn verkregen via Sigma-Aldrich, Merck en Toronto Research Chemicals. Aan de blanco's werden de interne standaarden atrazine-d5 (puur, Sigma), bentazon-d6 (puur, Sigma) en chlooroxuron (puur, sigma) toegevoegd. Om de modellen te testen werden tevens een oppervlaktewatermonster (Lekkanaal, 18 september 2020) en een spike-standaard gemeten.

2.4.2 LC-HRMS

De standaarden zijn gemeten met behulp van een Vanquish HPLC systeem (Thermo Fisher Scientific) gekoppeld aan een Tribrid Orbitrap Fusion massaspectrometer (Thermo Fisher Scientific). De analytische kolom XBridge BEH C18 XP (150 mm x 2,1 mm interne diameter, 2,5 µm deeltjes grootte, Waters, Etten-Leur, Nederland) werd voorafgegaan door een Phenomenex SecurityGuard Ultra column (UHPLC C18, 2,1 mm i.d., Phenomenex, Torrance, USA).

De mobiele fase bestond uit 0,05% mierenzuur (J.T. Baker, Avantor Performance Materials B.V., Deventer, Nederland) in ultrapuur water (LiChrosolv, LC-MS grade, Merck, Darmstadt, Duitsland) [v/v] (mobiele fase A) en 0,05% mierenzuur in acetonitril (J.T. Baker, ultra-gradient HPLC grade, Avantor Performance Materials B.V., Deventer, Nederland). Het injectievolume was 100 µL. De totale runtijd was 34 minuten: 0 tot 1 minuten isocratisch op 5% B; 1 tot 25 minuten lineair gradiënt tot 100% B; 25 tot 29 minuten isocratisch op 100% B; 29 tot 29,5 minuten lineair gradiënt tot 5% B; 29,5 tot 34 minuten isocratisch op 5% B.

Full scan hoge resolutie massaspectra werden opgenomen met de Orbitrap detector bij een resolutie van 120.000 FWHM van m/z 80 tot m/z 1300 gedurende de eerste 28 minuten van de LC analyse. De analyses werden in zowel positieve als negatieve ionisatiemodus uitgevoerd.

2.4.3 Bepaling RT en opname MS2 spectra

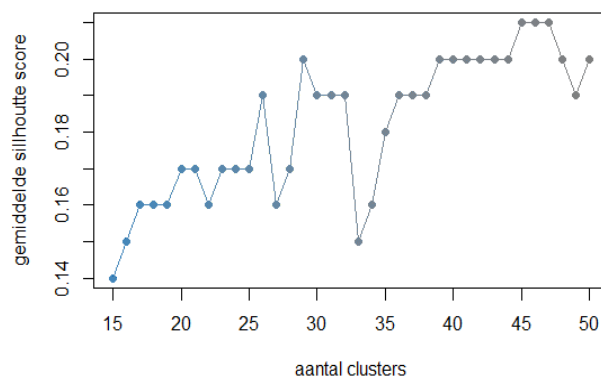
De verschillende standaarden werden bijeengevoegd in meerdere mixen in een concentratie van 20 µg/L in ultrapuur water. De monsters werden in triplo gemeten (gemiddelde relatieve standaarddeviatie = 0,2%). Vervolgens werd de retentietijd van elke stof bepaald uit de chromatogrammen aan de hand van de accurate massa.

2.5 Resultaten retentietijdvoorspelling

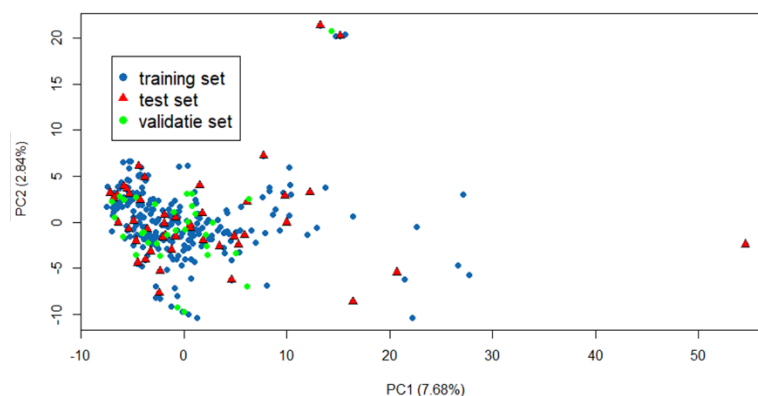
2.5.1 Verdeling van de data

Met behulp van PaDEL-Descriptor¹⁴ zijn 1444 descriptors berekend voor de 316 gemeten stoffen, na het opschonen van deze descriptors (verwijderen van descriptors met constante of ontbrekende waarden) zijn er 615 overgebleven. Met behulp van PCA (Principal Component Analysis, hoofdcomponentenanalyse) is een dimensiereductie uitgevoerd, de 44 principal components (PC's) die 50,1% van de variabiliteit verklaren zijn geselecteerd. Vervolgens zijn deze 44 PC's gebruikt om de stoffen in te delen in clusters met K-means clustering, het optimale aantal clusters is bepaald aan de hand van de silhouet-score (Figuur 7) (het aantal clusters met de hoogste gemiddelde silhouet-score is geselecteerd). De silhouet score geeft een indicatie van de mate waarin een stof juist is geclassificeerd, dus in hoeverre de stof op zijn eigen cluster lijkt en in hoeverre de stof van de andere clusters afwijkt. De silhouet score is weergegeven als een waarde tussen -1 en +1, waarbij een hoge score betekent

dat een stof goed bij het eigen cluster past, en niet goed bij de omliggende clusters. De hoogste gemiddelde silhouet score werd verkregen bij 45, 46 en 47 clusters, waarbij de verdeling in 46 clusters resulteerde in het kleinste aantal stoffen met een negatieve silhouet-score. Vervolgens werd uit elke cluster één stof overgezet naar de testset (14,6%), de overige stoffen werden verzameld in de trainingset. Deze trainingset wordt gebruikt om de modellen te berekenen. Vanuit deze trainingset is een willekeurige 10% van de stoffen overgezet naar de validatieset. Figuur 8 geeft de verdeling van de training-, test- en validatieset weer. Hieruit blijkt dat de testset goed verdeeld is over de gehele dataset, er valt slechts één stof (vancomycine, $C_{66}H_{75}Cl_2N_9O_{24}$) ver buiten de testset, op 57 van PC1 (x-as). Een verklaring hiervoor is dat dit molecuul een apart cluster vormt en er geen training stof aanwezig is.



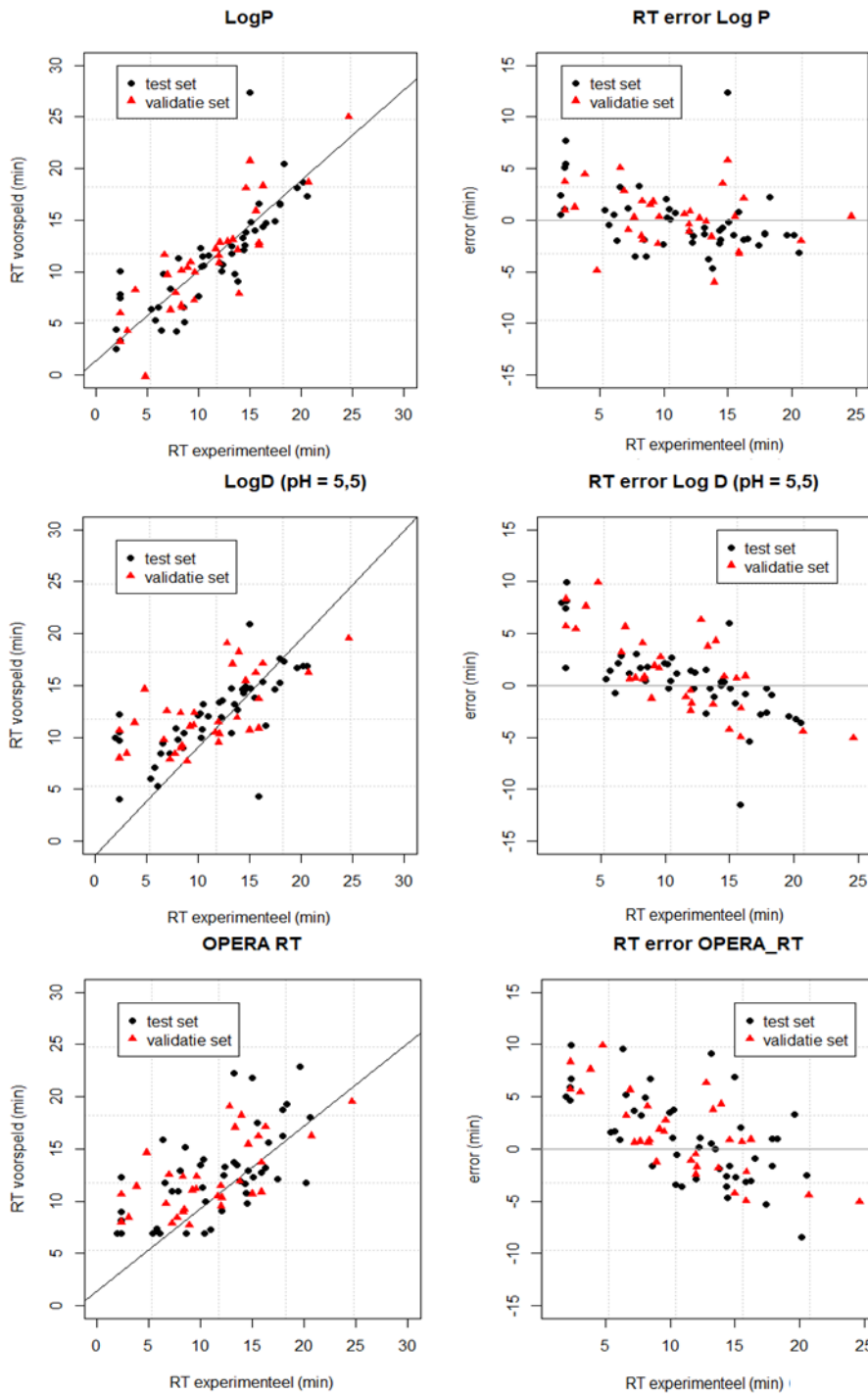
Figuur 7 - Aantal clusters uitgezet tegen de gemiddelde silhouet-score van de stoffen in de clusters.



Figuur 8 - PCA plot ter visualisatie van de verdeling van de training- en testset.

2.5.2 LogP, logD en OPERA-RT modellen (lineaire- en meervoudige-lineaire regressie)

De drie modellen met één variabele, logP (vergelijking 5), logD (vergelijking 6) en OPERA-RT (vergelijking 7) zijn gebruikt om de retentietijd te voorspellen, welke in de linker kolom in Figuur 9 uitgezet is tegen de experimentele retentietijd. De rechter kolom in Figuur 9 geeft voor elk model de fouten in de voorspellingen weer. Deze plots laten zien dat het logP model de betrouwbaarste voorspelling geeft ten opzichte van het logD en het OPERA-RT model. De spreiding rond de regressielijn is klein en bij de andere twee modellen zijn meer fouten groter dan 5 minuten dan bij het logP model. Aan de hand van de testresultaten blijkt dat het logD model bij een pH van 5,5 de nauwkeurigste voorspellingen geeft t.o.v. het logD model bij pH 7,4.



Figuur 9 - De experimentele retentietijd uitgezet tegen de voorspelde retentietijd (linker kolom) en de fouten in deze voorspelling (rechter kolom) per model.

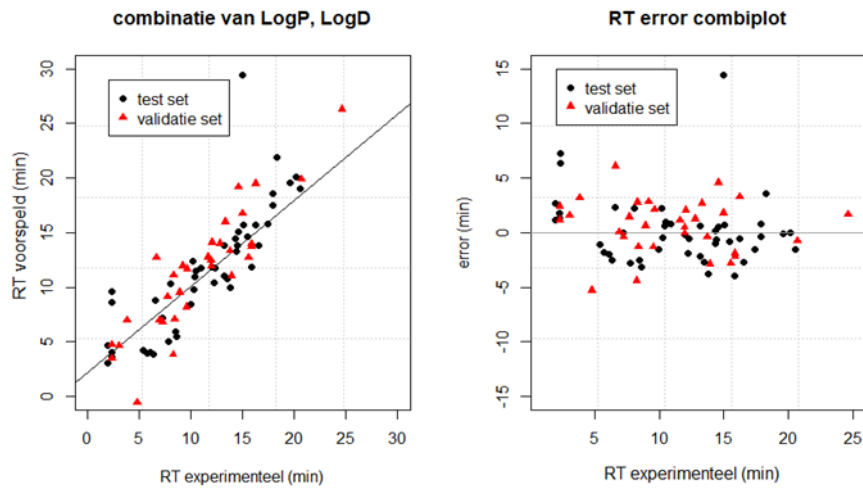
Het combinatiemodel bestaat uit de combinatie van de drie variabelen, samengesteld met behulp van multiple lineaire regressie. De variabele OPERA-RT heeft een p-waarde van 0,13 en is dus niet significant, om deze reden is deze variabele weggelaten uit het model (vergelijking 8). De resultaten van het combinatiemodel zijn uitgezet in Figuur 10. Uit de formule blijkt dat de logP parameter de grootste invloed heeft op de voorspelling, waardoor dezelfde uitbijter opnieuw aanwezig is.

$$Retentietijd_{logP\ model} = 2,75 \cdot LogP + 6,45 \tag{5}$$

$$Retentietijd_{logD\ model} = 1,41 \cdot LogD + 11,27 \tag{6}$$

$$\text{Retentietijd}_{\text{OPERA-RT model}} = 0,51 \cdot \text{OPERA} + 6,94 \quad (7)$$

$$\text{Retentietijd}_{\text{combinatiemodel}} = 2,28 \cdot \text{LogP} + 0,73 \cdot \text{LogD} + 7,04 \quad (8)$$



Figuur 10 - De experimentele retentietijd uitgezet tegen de door het combinatiemodel voorspelde retentietijd (links) en de fouten in deze voorspelling (rechts).

De modelprestaties worden met elkaar vergeleken in Tabel 1. Hieruit blijkt dat het logD en het OPERA-RT model niet voldoen aan de Golbraikh and Tropsha¹⁵ criteria die worden gesteld voor de validatieset. Tevens zijn er in de resultaten van de testset ook grote verschillen zichtbaar tussen deze twee modellen en het logP- en combinatiemodel. De resultaten van het logP en het combinatiemodel zijn vergelijkbaar, zowel voor de criteria van de validatieset als voor de test- en trainingset parameters. Het logP model heeft minder gegevens nodig, hierdoor is het eenvoudiger te gebruiken dan het combinatiemodel.

Tabel 1 - Vergelijking modelprestaties van het logP, logD, OPERA-RT en het combinatiemodel. De rode getallen voldoen niet aan de gestelde criteria.

	LogP model	LogD model	OPERA-RT model	Combinatiemodel
<i>Trainingset parameters</i>				
R ²	0,70	0,40	0,46	0,78
RMSE (min)	2,81	3,98	3,76	2,37
<i>Testset parameters</i>				
R ²	0,67	0,50	0,40	0,70
RMSE (min)	3,09	3,80	4,34	3,08
MAE (min)	2,23	2,60	3,53	1,94
MAPE (%)	30,58	57,15	61,88	33,25
Q ² _{loo}	0,66	0,48	0,32	0,66
<i>Validatieset criteria</i>				
Criterion I Q ² _{loo} > 0,5	0,72	0,42	0,42	0,8
Criterion II R ² > 0,6	0,75	0,42	0,42	0,79
Criterion III (R ² -R ₀ ²)/R ² < 0,1 en 0,85 < K' < 1,15	(0,75-0,73)/0,73 = 0,03 0,95	(0,42-0,42)/0,42 = 0,00 0,89	(0,42-0,42)/0,42 = 0,00 0,89	(0,79-0,79)/0,79 = 0,00 0,93
Criterion IV R ₀ ² -R ₀ ² < 0,3	0,73-0,73=0,00	-0,47-0,42=-0,89	-0,47-0,42=-0,89	0,79-0,79=0,00

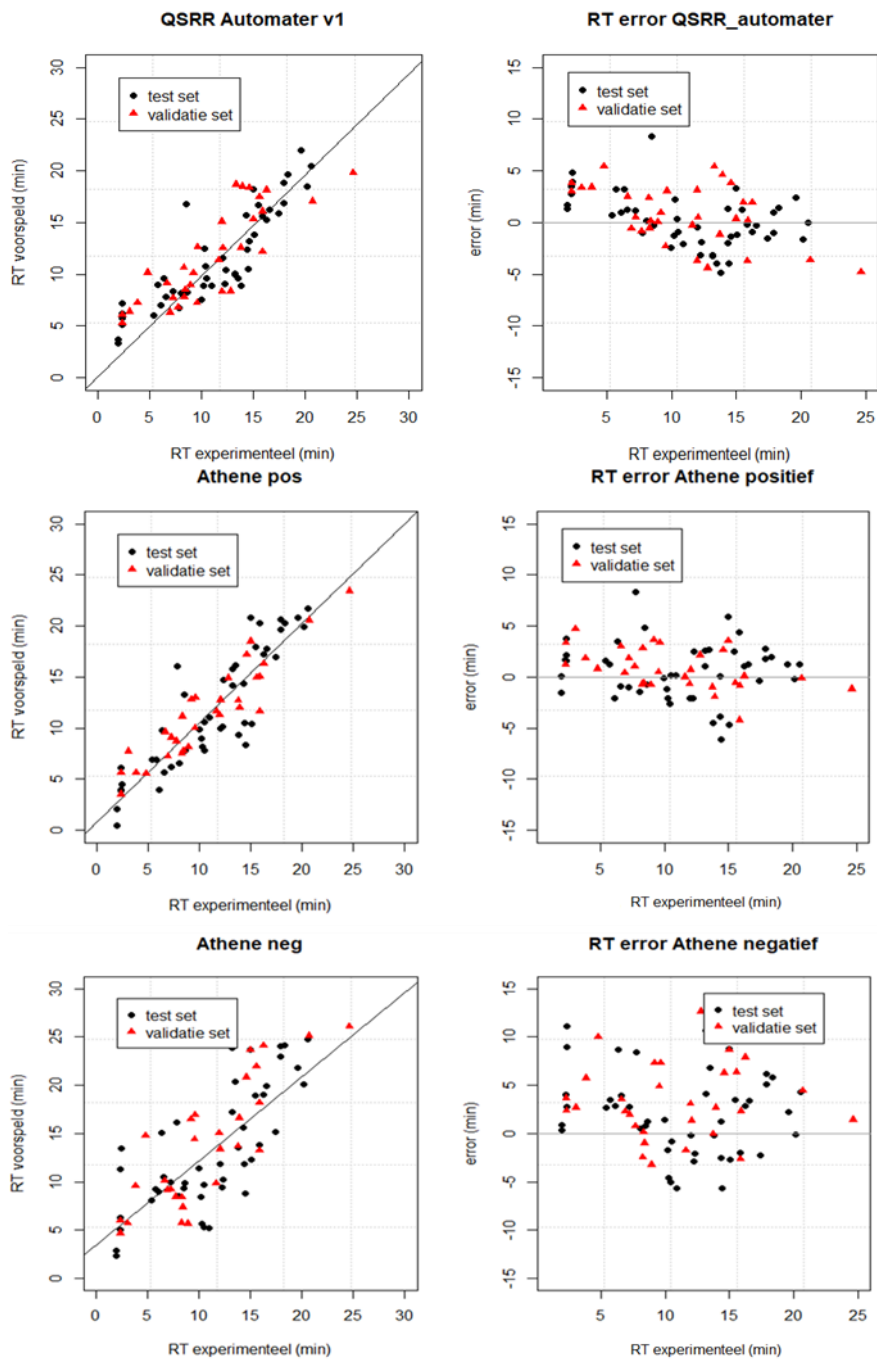
2.5.3 QSRR modellen

Met QSRR Automator werd een model op basis van 54 descriptors gegenereerd. De voorspelde retentietijd is in Figuur 11 uitgezet tegen de experimentele retentietijd, evenals de bijbehorende fouten. De meeste retentietijden worden binnen een bereik van ± 5 minuten voorspeld, dit lijkt nauwkeuriger dan het logP- en het combinatiemodel. Hier is tevens een uitbijter aanwezig (tetrapropylammonium, $C_{12}H_{28}N^+$), weliswaar een andere dan bij het logP- en het combinatiemodel. De modelprestaties worden weergegeven in Tabel 2. Wanneer deze vergeleken worden met het logP- en combinatiemodel blijkt dat het QSRR Automator model beter voorspelt.

In tegenstelling tot de andere modellen, is het QSRR model van Athene in twee verschillende uitvoeringen ontwikkeld, zowel voor positieve ionisatie als negatieve ionisatie. De modellen zijn gevormd op basis van een retentietijdindex met 18 standaarden. De voorspelde retentietijd is in Figuur 11 uitgezet tegen de experimentele retentietijd, evenals de bijbehorende fouten. Hieruit blijkt dat het model voor de positieve ionisatie minder spreiding heeft dan het model voor de negatieve ionisatie. Tevens voorspelt het positieve ionisatiemodel de meeste retentietijden binnen een bereik van ± 5 minuten. Uit de modelprestaties, weergegeven in Tabel 2, blijkt dat het model voor de negatieve ionisatie minder presteert dan het model voor de positieve ionisatie. Een mogelijke verklaring hiervoor is dat de in de praktijk toegepaste chromatografische condities in de negatieve modus sterk afwijken t.o.v. de condities die zijn toegepast voor de ontwikkeling van het QSRR model van Athene. Aan de hand van de parameters lijkt het QSRR Automator model iets betere voorspellingen te geven dan het positieve model van Athene.

Een kanttekening bij het model van Athene is dat bij drie stoffen de SMILES niet werden herkend en de retentietijd van deze stoffen niet voorspeld kon worden. Een voordeel is echter dat er geen grote trainingset (alleen de 36 standaarden) nodig is om het model te genereren omdat dit op basis van de RTI-standaarden gegenereerd wordt. Dit model is dus eenvoudiger te genereren voor laboratoria met verschillende chromatografische systemen, maar zou daardoor misschien ook minder nauwkeurig kunnen zijn indien het om een sterk afwijkend chromatografisch systeem gaat. Hier lijkt sprake van te zijn in het negatieve ionisatiemodel van Athene, mogelijk is er een grotere trainingset nodig om dit model te optimaliseren. Echter zal dit dan door de ontwikkelaars aangepast moeten worden, de grootte van de trainingset staat momenteel vast.

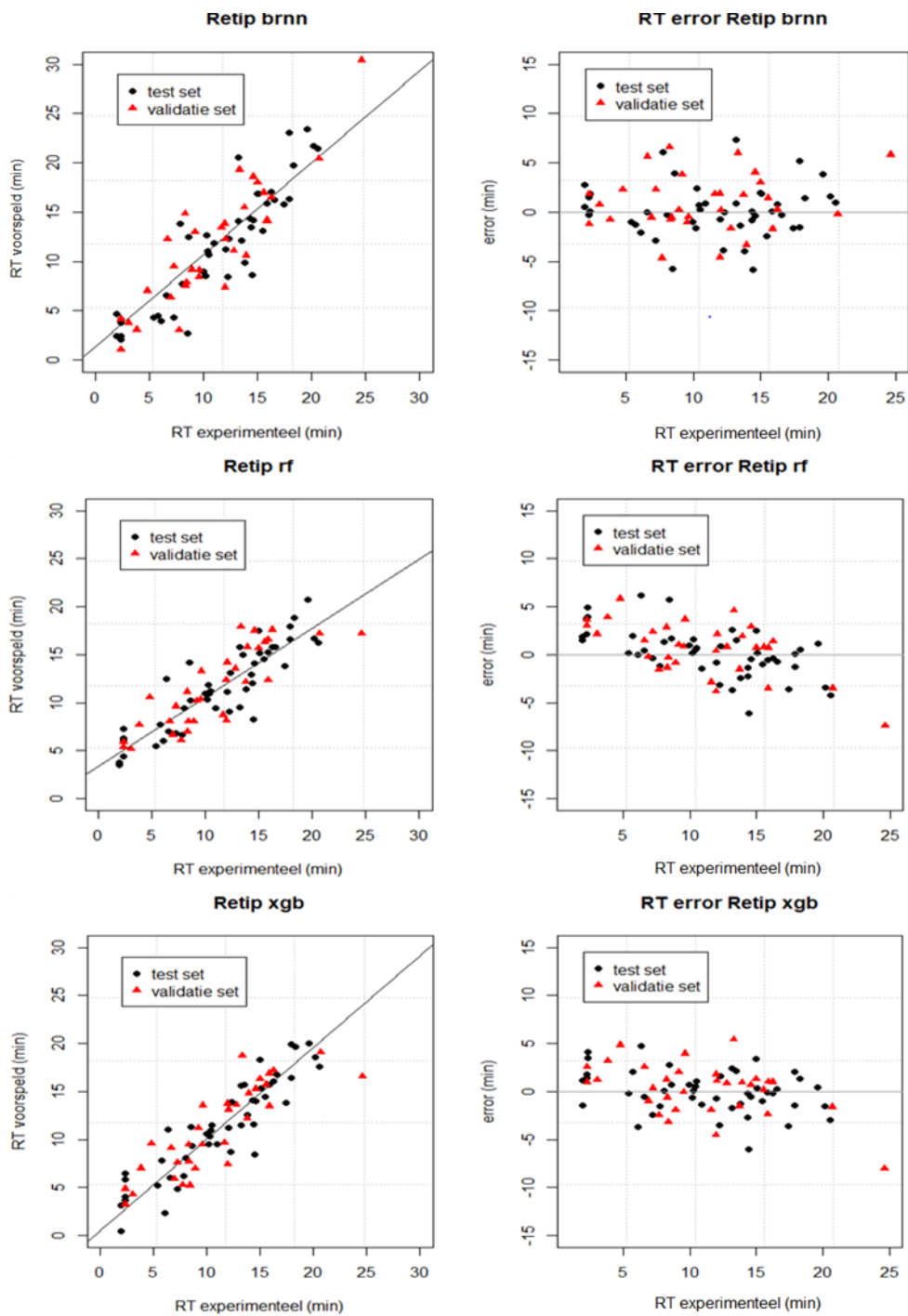
Met het R-pakket Retip zijn op basis van 145 descriptors drie modellen op drie verschillende manieren samengesteld: met Bayesian Regularized Neural Network (BRNN), een extreme Gradient Boosting voor drie algoritmen (XGB) en met behulp van Random Forest (RF). De voorspelde retentietijd is in Figuur 12 uitgezet tegen de experimentele retentietijd, evenals de bijbehorende fouten. Deze resultaten laten zien dat de spreiding klein is, waarbij het XGB model de nauwkeurigste voorspelling geeft, maar wel een uitbijter aanwezig is op 24,6 minuten (eicosapentaenoic acid, $C_{20}H_{30}O_2$). Het BRNN model laat de grootste spreiding zien. De modelprestaties worden weergegeven in Tabel 3. Op basis van deze waarden blijkt dat het Retip-XGB model de nauwkeurigste voorspelling geeft. De waarden van RMSE, MAE en MAPE liggen lager en de Q^2_{LOO} waarde ligt hoger.



Figuur 11 - De experimentele retentietijd uitgezet tegen de voorspelde retentietijd (linker kolom) en de fouten in deze voorspelling (rechter kolom) per model.

Tabel 2 - Vergelijking modelprestaties van het QSRR Automator model en de QSRR modellen van Athene. Het rode getal voldoet niet aan de gestelde criteria.

	QSRR Automator	QSRR Athene (negatief)	QSRR Athene (positief)
<i>Trainingset parameters</i>			
R ²	0,97	n.v.t.	n.v.t.
RMSE (min)	0,99	n.v.t.	n.v.t.
<i>Testset parameters</i>			
R ²	0,77	0,55	0,78
RMSE (min)	2,56	4,69	2,83
MAE (min)	2,02	3,74	2,21
MAPE (%)	32,52	55,65	28,51
Q ² _{loo}	0,76	0,22	0,72
<i>Validatieset criteria</i>			
Criterium I Q ² _{loo} > 0,5	0,72	0,07	0,83
Criterium II R ² > 0,6	0,67	0,67	0,86
Criterium III (R ² -R ₀ ²)/R ² < 0,1 en 0,85 < K' < 1,15	(0,67-0,67)/0,67 = 0,00 1,01	(0,67-0,65)/0,67 = 0,03 0,74	(0,86-0,85)/0,85 = 0,01 0,94
Criterium IV R ₀ ² -R ₀ ² < 0,3	0,57-0,67=0,10	0,65-0,63=0,02	-0,47-0,42=0,06



Figuur 12 - De experimentele retentietijd uitgezet tegen de voorspelde retentietijd (linker kolom) en de fouten in deze voorspelling (rechter kolom) per model.

Tabel 3 - Vergelijking modelprestaties van de verschillende Retip modellen.

	Retip XGB	Retip RF	Retip BRNN
<i>Trainingset parameters</i>			
R ²	0,95	0,79	0,99
RMSE (min)	1,17	1,11	0,56
<i>Testset parameters</i>			
R ²	0,83	0,78	0,59
RMSE (min)	2,18	2,51	4,2
MAE (min)	1,7	1,9	2,35
MAPE (%)	26,07	30,94	30,01
Q ² _{LOO}	0,82	0,77	0,63
<i>Validatieset criteria</i>			
Criterium I Q ² _{LOO} > 0,5	0,72	0,68	0,66
Criterium II R ² > 0,6	0,73	0,70	0,79
Criterium III (R ² -R ₀ ²)/R ² < 0,1 en 0,85 < K' < 1,15	(0,73-0,73)/0,73 = 0,00 0,98	(0,70-0,70)/0,70 = 0,00 0,96	(0,79-0,75)/0,79 = 0,05 0,89
Criterium IV R ₀ ² -R ₀ ² < 0,3	0,73-0,65=0,08	0,70-0,51=0,19	0,79-0,75=0,04

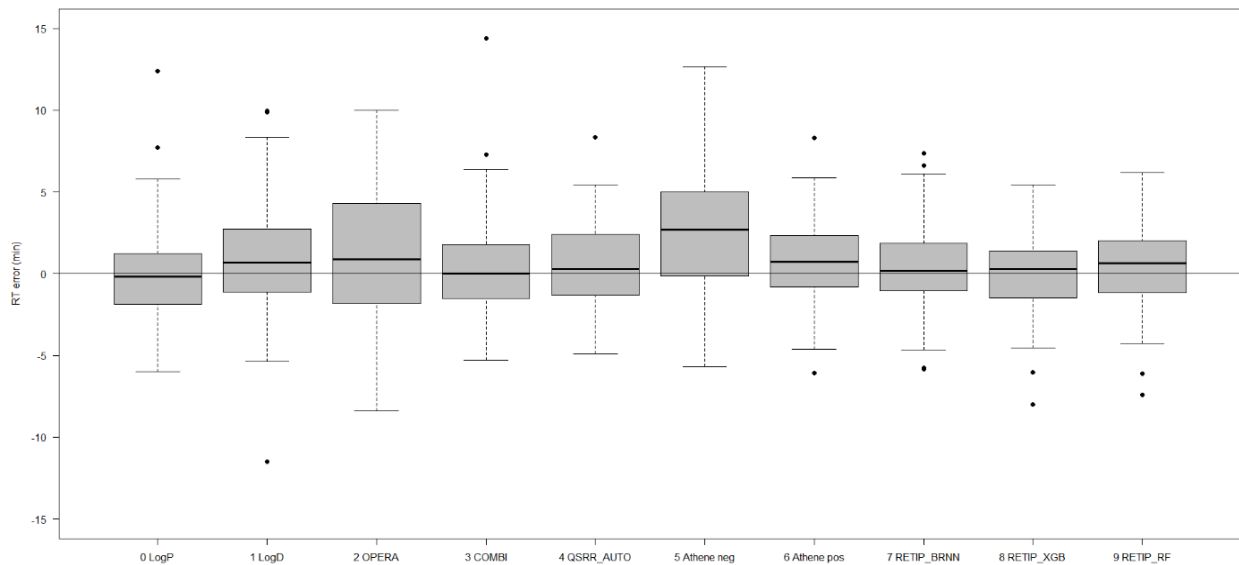
Tijdens de bepaling van de theoretische parameters en criteria zijn de retentietijden van 78 stoffen (46 uit de testset en 32 uit de validatieset, op basis van een trainingset van 238 stoffen) voorspeld. Met behulp van het retentietijdgebied kan worden vergeleken hoeveel stoffen er door ieder model binnen dit gebied correct worden voorspeld. Dit gebied wordt aangegeven als percentage van de retentietijd, zowel boven als onder de retentietijd. Deze resultaten zijn weergegeven in Tabel 4.

Tabel 4 - Aantal componenten en als percentage van het totaal dat binnen het gegeven retentietijdgebied voorspeld wordt.

Model	Bereik ±5% (±1,45min)	Bereik ±10% (±2,9min)	Bereik ±15% (±4,35min)	Bereik ±20% (±5,8min)
LogP	37 = 47%	58 = 74%	68 = 87%	75 = 96%
LogD	31 = 40%	51 = 65%	60 = 77%	67 = 86%
OPERA-RT	20 = 26%	38 = 49%	53 = 68%	66 = 85%
Combinatie	35 = 45%	65 = 84%	71 = 91%	74 = 95%
QSRR Automator	35 = 45%	50 = 64%	70 = 90%	77 = 99%
QSRR Athene positief*	34 = 45%	58 = 74%	67 = 89%	73 = 97%
QSRR Athene negatief*	17 = 23%	38 = 51%	50 = 67%	59 = 79%
Retip BRNN	36 = 46%	58 = 74%	66 = 85%	72 = 92%
Retip RF	36 = 46%	56 = 76%	71 = 91%	74 = 95%
Retip XGB	38 = 49%	62 = 79%	72 = 92%	76 = 97%

* Deze modellen hadden 3 stoffen minder.

Uit de resultaten blijkt dat het QSRR Automator de meeste stoffen correct voorspelt in een gebied van ±20%. Wanneer dit gebied verkleind wordt, geven het Retip-RF, Retip-XGB en het combinatiemodel de beste resultaten. Dit wordt ondersteund door Figuur 13, waar de spreiding van de fouten van alle modellen wordt weergegeven. De figuur laat zien dat het Retip-XGB model de kleinste spreiding heeft, maar meer uitbijters dan het QSRR Automator model.



Figuur 13 - Spreiding van de fouten van alle modellen (test- en validatiedata). De boxplots geven de mediaan (centrale streep), het 25 en 75 percentiel (box) en de minimale en maximale waarden weer, de uitbijters worden weergegeven met de zwarte stippen.

2.6 Modelprestaties in de praktijk

Vervolgens werden de modellen toegepast in de reguliere NTS workflow om zo het aantal kandidaatstoffen voor de onbekende features te verlagen. Een spikestandaard werd gemeten en de standaarden werden geïdentificeerd met behulp van de accurate massa, retentietijd en intensiteit. De features werden vergeleken met twee databases: ChemSpider (> 67 miljoen stoffen) en de NORMAN Substance Database (37.028 stoffen). Met behulp van deze databases konden zestien van de twintig standaarden worden geïdentificeerd. In Tabel 5 is de afname van het aantal ChemSpider kandidaten per model weergegeven, met een retentietijdbereik van ± 5 minuten. Tabel 6 geeft hetzelfde weer maar dan voor de NORMAN Substance Database kandidaten.

Tabel 5 - Reductie van het aantal kandidaatstoffen bij een bereik van ± 5 minuten, de kandidaatstoffen zijn verkregen met de ChemSpider database. In de gekleurde vlakken staat de hoeveelheid kandidaten die overblijft na filteren. Rood vlak: de juiste kandidaat is weggefilterd door het model. Oranje vlak: de juiste kandidaat was niet aanwezig in de voorspelling. Groen vlak: de juiste kandidaat is na de reductiestap nog steeds aanwezig. Wit vlak: er kon geen voorspelling worden gedaan.

Feature	Aantal kandidaten	Retip-BRNN	Retip-XGB	Retip-RF	QSRR Automator	Athene positief	LogP	Combinatiemodel
1	20	6	7	7	4	7	0	0
2	20	17	16	15	14	13	5	12
3	19	14	17	19	18	12	6	5
4	19	11	12	13	11		9	
5	20	9	4	3	1	5	2	2
6	20						1	2
7	20	1	20	20	20	20	0	0
8	12	8	8	8			6	
9	7	0	7	7	7	7	6	5
10	20	11	9	17	9	11	0	4
11	20	13	9	6	3	5	19	18
12	3	3	3	3	3		3	
13	12	9	9	12	12	12	12	12
14	20	17	17	17	20		15	
15	20	7	7	7	7	7	7	
16	20	20	20	20	20	20	20	20

Wanneer hetzelfde wordt gedaan met een bereik van ± 3 minuten (data niet weergegeven) blijkt dat het vaker voorkomt dat de 'echte' stof wordt weg gefilterd. Voor feature 6 (strychnine, $C_{21}H_{22}N_2O_2$) werd een SMILES gegenereerd die niet verwerkt kon worden door de modellen, behalve door het logP en combinatiemodel omdat er wel een logP waarde gegenereerd kon worden. De Retip modellen kunnen geen voorspelling geven wanneer er metalen aanwezig zijn in het molecuul, of wanneer het molecuul in de zoutvorm wordt ingeladen. QSRR-automator kan in deze gevallen echter wel een voorspelling geven. Een nadeel van QSRR-automator is dat slechts een kleine selectie van descriptoren wordt meegenomen terwijl de Retip modellen er meer meenemen, waardoor de Retip modellen ook stoffen kunnen voorspellen die afwijken van de trainingset.

Tabel 6 - Reductie van het aantal kandidaatstoffen bij een bereik van ± 5 minuten, de kandidaatstoffen zijn verkregen met de NORMAN Substance database. In de gekleurde vlakken staat de hoeveelheid kandidaten die overblijft na filteren. Rood vlak: de juiste kandidaat is weggefilterd door het model. Oranje vlak: de juiste kandidaat was niet aanwezig in de voorspelling. Groen vlak: de juiste kandidaat is na de reductiestap nog steeds aanwezig.

Feature	Aantal kandidaten	Retip-BRNN	Retip-XGB	Retip-RF	QSRR Automator	Athene positief	LogP	Combinatiemodel
1	8	3	3	3	3	3	1	2
2	28	14	13	13	11	12	1	4
3	2	2	2	2	2	2	2	2
4	3	1	3	3	3	3	2	3
5	1	1	1	1	1	1	1	1
6	5	4	2	4	5	3	0	0
7	1	1	1	1	1	1	1	1
8	2	1	1	1	1	1	1	1
9	3	0	1	1	1	0	0	0
10	1	1	1	1	1	1	1	1
11	1	1	1	0	0	1	1	1
12	1	1	1	1	1	1	0	0
13	5	5	5	5	5	5	5	5
14	13	5	4	4	4	5	7	7
15	6	4	4	4	4	4	4	4
16	11	11	11	11	11	11	11	11

Al met al lijken het Retip-XGB model, Retip-RF model en het QSRR-Automator model de beste resultaten te geven. Deze drie modellen zijn toegepast voor de reductie van de kandidaatstoffen voor de features in het oppervlaktewater monster. De gedetecteerde features werden geannoteerd met behulp van twee verschillende databases: de Sjerps database (5.600 stoffen) en de SusDat database (37.028 stoffen). Voor elk feature wordt zo een aantal kandidaatstoffen gegenereerd. Vervolgens werden de retentietijden van de kandidaten voorspeld en gefilterd met een range van ± 5 minuten. In Tabel 7 is de gemiddelde reductie van het aantal kandidaten na het filteren weergegeven.

Tabel 7 - Reductie van kandidaten na filteren. Het betreft hier MS1-data in de SusDat en Sjerps databases, voor het oppervlaktewatermonster. Filtering vindt plaats door middel van een bereik van ± 5 minuten.

	Retip-XGB	Retip-RF	QSRR Automator
<i>Positieve ionisatie</i>			
Reductie van SusDat kandidaten na filteren (%)	29,3	28,9	23,9
Reductie van Sjerps kandidaten na filteren (%)	23,3	22,1	23,4
<i>Negatieve ionisatie</i>			
Reductie van SusDat kandidaten na filteren (%)	38,8	36,3	43,3
Reductie van Sjerps kandidaten na filteren (%)	37,5	37,5	56,3

Naast suspect screening met behulp van de SusDat en de Sjerps databases is er ook een suspect screening uitgevoerd met behulp van de *in silico* fragmentatie software MetFrag¹⁶, dit is toegepast op features met een MS2 spectrum. De modellen zijn toegepast op de 200 kandidaten met de hoogste MetFrag score. Het gemiddelde aantal gereduceerde kandidaten is weergegeven in Tabel 8. Voor het Retip-XGB model is een reductie van 1,6% zichtbaar

voor de negatieve ionisatiemodus t.o.v. hogere percentages voor de andere twee modellen. Dit komt doordat er weinig kandidaatstoffen (soms slechts één stof) gevonden zijn voor de features die gemeten zijn in de negatieve ionisatiemodus, dit leidt tot grote verschillen in gereduceerde kandidaten.

Tabel 8 - Reductie van kandidaten na filteren van MS2 data door middel van de nauwkeurigste modellen. Filtering vindt plaats door middel van een bereik van ± 5 minuten.

	Retip-XGB	Retip-RF	QSRR Automator
Reductie van kandidaten na filtering in positieve ionisatiemodus (%)	42,1%	36,2%	37,4%
Reductie van kandidaten na filtering in negatieve ionisatiemodus (%)	1,6%	22,5%	24,6%

2.7 Discussie van de modelprestaties

De test- en validatieset laten zien dat het Retip-XGB model de nauwkeurigste voorspelling geeft: 48,7% van alle stoffen uit de test- en validatieset werd binnen een bereik van 5% van de totale retentietijd correct voorspeld. Dit model is ook in andere studies getest en daar zijn vergelijkbare scores verkregen (zie Tabel 9). De ontwikkelaars van Retip-XGB rapporteren een kleinere fout (0,48 minuten voor de testset van Bonini, Kind, Tsugawa, Barupal and Fiehn ⁸ t.o.v. 1,17 minuten in dit onderzoek). Dit kan worden verklaard doordat de ontwikkelaars een grotere trainingset gebruikten op basis van 398 stoffen, de grootte van de trainingset heeft invloed op de nauwkeurigheid van de voorspelling¹⁷.

Tabel 9 - Vergelijking prestaties Retip-XGB model met waarden uit de literatuur.

	R ²	Q ² _{loo}	Criterium III voor validatieset
Retip-XGB model in dit onderzoek	0,83	0,82	0,00 & 0,98
Aalizadeh, Thomaidis, Bletsou and Gago-Ferrero ¹⁰	0,84	0,81	0,09 & 1,00
McEachran, Mansouri, Newton, Beverly, Sobus and Williams ¹¹	0,83	0,83	NA
Bonini, Kind, Tsugawa, Barupal and Fiehn ⁸	0,90	0,80	NA

De ontwikkelde modellen kunnen worden ingezet om het aantal kandidaatstoffen te reduceren binnen NTS, echter is deze afname minder groot dan verwacht. De test met MS2-data van MetFrag bij de positieve ionisatie heeft de meeste kandidaatstoffen opgeleverd en deze resultaten geven de beste indicatie van de modelprestaties bij het filteren. Het Retip-XGB model geeft de nauwkeurigste voorspellingen. QSRR Automator model is het nauwkeurigste model na de Retip modellen, hier treden geen problemen op wanneer het molecuul metalen bevat of wanneer de SMILES in de zoutvorm worden aangedragen. Echter bevat dit model minder descriptoren omdat deze door backward elimination verwijderd zijn. Backward elimination is een proces waarbij het model ontwikkeld wordt met alle descriptoren, en deze één voor één verwijderd worden als er geen significant verschil zichtbaar is (wanneer dit wel het geval is wordt de descriptor niet verwijderd). Een gevolg hiervan is dat het model een slechte tot geen voorspelling geeft voor moleculen die sterk afwijken van de moleculen uit de trainingset. Uit de test met het oppervlaktewatermonster is dan ook gebleken dat de nadelen van beide modellen ertoe leiden dat ongeveer hetzelfde aantal kandidaten wordt gereduceerd (Tabel 8).

Het logD model lijkt de minst nauwkeurige voorspelling te geven, de resultaten suggereren dat een logD bij een lagere pH kan leiden tot een betere voorspelling. Nu zijn alleen pH 7,4 en pH 5,5 getest, wellicht geeft een logD bij een pH die dichterbij de zuurgraad van het eluens (pH 2,9) ligt, nauwkeurigere voorspellingen. De lage correlatie tussen de door OPERA voorspelde retentietijd en de experimentele retentietijd is te verklaren door de vele verschillen in de chromatografische systemen. Dit komt ook naar voren bij de voorspellingen van het QSRR model

van Athene voor negatieve ionisatie, de chromatografie waarop het model is ontwikkeld wijkt waarschijnlijk te veel af van de toegepaste chromatografie in het huidige onderzoek.

2.8 Toepassing op NTS dataset

Een triplo meting (positieve modus) van een onbekend monster (met monstercode 'S1a') vanuit het NORMAN ringonderzoek werd geanalyseerd met Compound Discoverer 3.2 (Thermo Fisher Scientific, San Jose), volgens de standaard workflow voor de analyse van NTS data. Het maximaal aantal kandidaten vanuit ChemSpider werd op 100 gesteld. Vervolgens zijn de 50 bekende (gespikede) stoffen teruggezocht in de resultaten en zijn de bijbehorende features genoteerd. Voor al deze features is uitgezocht of de ChemSpider resultaten de werkelijke stof bevatten, en of de stof niet al gebruikt is voor de ontwikkeling van de modellen. Uiteindelijk bleven er voor de positieve modus 19 features over waarvan de identiteit bekend is, de goede stof tussen de ChemSpider resultaten zit en de stof niet gebruikt is voor de ontwikkeling van het model. Deze zijn weergegeven in Tabel 10.

Tabel 10 - Modelprestaties voor 19 features met bekende identiteit. Groen betekent dat de echte stof nog steeds tussen de kandidaten zit.

Naam en Feature ID	n	Voorspelde RT van juiste stof				Aantal features binnen +/- 3 min.			
		BRNN (Δ)	RF (Δ)	XGB (Δ)	QSRR automator (Δ)	BRNN	RF	XGB	QSRR automator
Atrazine-desethyl-desisopropyl 145.01526 / 2.532	6	4,85 (2,32)	6,20 (3,67)	6,68 (4,15)	5,21 (2,68)	4	1	1	2
Omethoate 213.02214 / 4.096	4	7,45 (3,35)	5,67 (1,57)	6,26 (2,16)	8,62 (4,52)	3	4	3	3
Benzotriazole-5-carboxylic acid 163.038 / 6.334	100	7,77 (1,44)	7,22 (0,89)	6,39 (0,06)	7,76 (1,43)	57	73	83	67
Thiabendazole 201.036 / 6.975	57	11,01 (4,04)	10,74 (3,76)	9,86 (2,89)	14,45 (7,47)	4	18	17	8
Atrazine-desisopropyl 173.04674 / 7.687	7	7,27 (0,42)	7,87 (0,18)	7,69 (0,003)	7,66 (0,023)	3	7	4	7
5-chlorobenzotriazole 153.00928 / 11.308	71	10,73 (0,58)	11,15 (0,16)	12,2 (0,89)	10,71 (0,6)	42	43	40	29
Carbamazepine-10,11-epoxide 252.08952 / 11.419	100	10,77 (0,65)	8,76 (2,66)	8,4 (3,02)	14,2 (2,78)	63	89	86	45
Benzothiazole 135.01398 / 12.379	17	13,05 (0,67)	10,66 (1,72)	11,32 (1,06)	12,81 (0,43)	12	14	13	13
Ketoconazole 530.14865 / 13.670	5	18,14 (4,47)	13,58 (0,09)	14,34 (0,67)	17,19 (3,52)	0	5	4	0
Phenazine 180.06859 / 14.298	54	10,31 (3,99)	9,51 (4,79)	10,05 (4,25)	14,89 (0,6)	25	24	27	48
Climbazole 292.09767 / 14.356	56	14,15 (0,21)	16,85 (2,49)	17,85 (3,49)	17,51 (3,15)	28	29	24	41
Reserpine 608.27281 / 14.825	3	16,46 (1,63)	13,28 (1,54)	14,43 (0,39)	17,38 (2,55)	3	3	3	3
Clotrimazole 344.10769 / 15.158	26	16,78 (1,62)	17,75 (2,59)	18,66 (3,50)	18,01 (2,86)	16	15	10	* 17
Irgarol 253.13597 / 15.2	3	17,73 (2,53)	16,01 (0,81)	16,92 (1,72)	14,29 (0,91)	1	3	3	3
Methidathion 301.96181 / 17.166	3	22,02 (4,85)	16,56 (0,61)	16,52 (0,65)	10,55 (6,61)	0	2	2	1
2-(methylthio)benzothiazole 181.0018 / 17.387	22	15,77 (1,62)	13,8 (3,59)	13,81 (3,58)	15,86 (1,52)	17	1	4	13
Simvastatin 418.27161 / 22.23	10	26,28 (4,05)	19,53 (2,7)	20,77 (1,46)	21,82 (0,41)	4	2	6	8
Sudan I 248.09476 / 23.138	95	22,72 (0,42)	15,51 (7,63)	16,27 (6,87)	17,97 (5,17)	12	0	0	0
Chlorpyrifos 348.92627 / 23.309	2	22,49 (0,82)	20,88 (2,43)	21,5 (1,81)	21,82 (1,49)	2	1	1	2
Gemiddeld % kandidaten dat overblijft na filteren						50,7	60,4	57,2	60,9
% van de features waarbij de juiste stof is weg gefilterd						31,6	26,3	36,8	31,5

* Totaal aantal kandidaten voor clotrimazole in QSRR Automator is 24, twee stoffen werden niet geaccepteerd door de software, deze zijn verwijderd.

Uit de resultaten in Tabel 10 blijkt dat de geteste modellen vergelijkbare prestaties vertonen. Met een range van +/- 3 minuten blijft 50-61% van de kandidaten nog over waarbij in 26-37% van de gevallen de juiste stof is verwijderd. Dit percentage zal toenemen als de range verkleind wordt, en afnemen als de range vergroot wordt.

Echter is een grotere range (+/- 5 minuten of meer) niet wenselijk omdat het aantal kandidaten dat overblijft ook groter wordt, m.a.w. er worden minder kandidaten weg gefilterd.

2.9 Gebruiksgemak en koppeling met patRoon

QSRR Automator is een eenvoudig te bedienen, losstaand, programma met een gebruikersinterface waardoor er geen programmeerkennis is vereist. De resultaten worden per feature opgeslagen in een Excelbestand en zijn op die manier makkelijk te verwerken. Dit loopt echter snel op wanneer grotere aantallen features bekeken worden met als resultaat dat het een tijdrovend proces wordt. De verschillende Retip modellen worden toegepast in een R-script, wat enige kennis van deze programmeertaal vereist. Het voordeel van deze software is dat grotere datasets op deze manier eenvoudiger te verwerken dan met QSRR Automator, ook omdat de Retip modellen meegenomen zouden kunnen worden in bestaande workflows (bijvoorbeeld op basis van patRoon¹⁸).

Beide modellen kunnen worden geïmplementeerd in patRoon, Retip zal het eenvoudigst te koppelen zijn met patRoon, De modellen zullen eerst door de gebruiker opgebouwd moeten worden door een voldoende grote trainingset (d.w.z. referentiestandaarden) te meten. En elke andere chromatografische methode zal een nieuw model vereisen.

2.10 Conclusie retentietijdvoorspellende modellen

De modellen Retip-XGB en QSRR Automator komen het beste naar voren uit dit onderzoek, beide modellen laten vergelijkbare resultaten zien. Echter wordt er niet voldoende winst behaald met het filteren van de kandidaten en worden nog te vaak de juiste stoffen verwijderd bij een range van +/- 3 minuten. Er worden betere resultaten behaald bij een range van +/- 5 minuten, echter worden er dan te weinig kandidaten gefilterd. Dit is in overeenstemming met de literatuur, het is nog steeds een uitdaging om deze range zo klein mogelijk te maken.⁸ De verwachting is dat de modellen geoptimaliseerd kunnen worden door een grotere set aan relevante stoffen te meten en te gebruiken voor de ontwikkeling van de modellen. Hoe complexer de stoffen zijn, des te groter moet de training set zijn om de verschillen te dekken.⁹ Dit zal dan voor elk uniek chromatografisch systeem en laboratorium uitgevoerd moeten worden, wat zeer tijdrovend is. Op dit moment lijkt het veelvuldig inzetten van retentietijdvoorspellende modellen nog niet van toegevoegde waarde voor de reductie van het aantal kandidaten voor features gevormd tijdens NTS. Het is beperkt bruikbaar, voor een aantal stoffen werkt de reductie van de kandidaten goed met een voldoende grote range, maar dit is niet voor alle stoffen het geval. Mogelijk is het waardevol om een nieuwe verkenning uit te voeren zodra er nieuwe retentietijdvoorspellende modellen ontworpen zijn aangezien deze ontwikkelingen door blijven gaan. De verwachting is dat deze steeds beter worden door onder andere de toenemende beschikbaarheid van trainingdata.¹⁹

3 Semikwantitatieve NTS

De ionisatie-efficiëntie van een stof bepaalt hoe succesvol een stof geïoniseerd kan worden. Een goed ioniseerbare stof (hoge ionisatie-efficiëntie) zal bij dezelfde concentratie een hoger signaal in het massaspectrum geven dan een slecht ioniseerbare stof (lage ionisatie-efficiëntie). Ionisatie-efficiëntie (IE) kan worden gebruikt om op basis van de oppervlakte of intensiteit van een piek, de concentratie van de bijbehorende stof te voorspellen. Het is vergelijkbaar met de responsfactor (RF) welke de verhouding tussen de oppervlakte van een piek en de concentratie weergeeft.

$$\text{Responsfactor (RF)} = \frac{\text{piekoppervlakte}}{\text{concentratie}}$$

Deze ionisatie-efficiëntie kan relaties hebben met verschillende andere parameters, zoals de zuur- en baseconstantes (pKa en pKb).^{5, 20-22} Krueve et al.²² observeerden een correlatie tussen log IE (ionisatie-efficiëntie) en pKa, waarbij sterke zuren (lage pKa waarden) de neiging tot hogere IE hebben. De mate van delocalisatie van lading in een anion (bij negatieve ionisatie) kan worden gekwantificeerd aan de hand van de WAPS (weighted average positive sigma) parameter.²² Deze WAPS parameter kan worden berekend met de software COSMO-RS. Des te lager de WAPS waarde, des te meer is de lading in het anion gedelokaliseerd. Krueve et al.²² observeerde hogere IE voor ionen met een efficiëntere delocalisatie van de lading (lagere WAPS). Een ander voorbeeld is de hydrofobiciteit van een stof, naarmate een stof hydrofober is zal deze beter te ioniseren zijn.²³ Dit komt onder andere doordat ze makkelijker accumuleren aan de oppervlakte van geladen druppels door de grotere oppervlakteactiviteit.²⁴ Voor de mobiele fase geldt dat de hoeveelheid organisch oplosmiddel, pH van de buffer en de compositie van de buffer allemaal van invloed zijn op de IE in ESI-MS.²⁵⁻²⁷ In het algemeen geldt dat een grotere hoeveelheid organisch oplosmiddel en een lagere pH in hogere IE resulteren, en daarmee een hogere respons in positieve ESI/MS.

Malm et al.²⁸ hebben een gids en overzicht voor semikwantitatieve NTS met LC/ESI/HRMS gepubliceerd, ze beschrijven een aantal regels met betrekking tot de ionisatie-efficiëntie (IE) voor verschillende stoffen. Over het algemeen hebben stoffen die sterker hydrofoob zijn een hogere IE, en in positieve modus zijn sterke basen in het voordeel (de beste IE wordt behaald in zuur milieu omdat ze dan sterk geprotoneerd worden²⁹), terwijl in negatieve modus de sterke zuren een hogere IE (welke het hoogste is in basisch milieu) hebben. Daarnaast is gebleken dat een grotere hoeveelheid organisch oplosmiddel ervoor zorgt dat de IE toeneemt, dit betekent dat stoffen die later elueren in reversed-phase chromatografie vaker hogere IE hebben. Dit wordt onder andere veroorzaakt door de betere verdamping van de mobiele fase druppels gedurende het ionisatieproces.²⁹ Ten slotte geven ze een overzicht van de verschillende stappen in NTS en het effect op de nauwkeurigheid van de semikwantificatie. Deze maatregelen lopen uiteen van zo min mogelijk monstervoorbewerking tot een aanbeveling van interne standaarden en aanbevolen scan ranges.²⁸

Ligand et al.³⁰ onderzocht in hoeverre IE-waarden die op het ene instrument bepaald zijn, overgedragen kunnen worden naar andere instrumenten. Hieruit is gebleken dat met behulp van 5-6 gangbare stoffen de IE waarden overgedragen kunnen worden. Dit principe wordt gebruikt voor de ontwikkeling van diverse voorspellende modellen, zie Paragraaf 3.1.3.

3.1 Methoden om ionisatie-efficiëntie te voorspellen

Semikwantificatie van features in non-target screening is in opkomst maar wordt nog niet op grote schaal of veelvuldig toegepast.³¹ Er zijn reeds verschillende methoden ontwikkeld om de IE van (onbekende) stoffen te voorspellen. Deze lopen uiteen van relatief eenvoudig (op basis van suspects) tot zeer complex (aan de hand van allerlei machine learning algoritmes). Een aantal van deze methoden wordt hieronder besproken, aan de hand van de kritische review van Malm et al.²⁸

3.1.1 Vergelijkbare stoffen

Deze methode vereist een suspect, dus een verwachte kandidaat van een onbekende piek, en is daarom niet toepasbaar bij 'echte' NTS data. Er wordt gebruik gemaakt van de responsfactor (RF) van een vergelijkbare stof.²⁸ De onderliggende aanname van deze methode is dat vergelijkbare stoffen ook een vergelijkbare ionisatie-efficiëntie hebben. Een vergelijkbare stof is een stof met overeenkomsten in de structuur, des te minder overeenkomsten, des te minder betrouwbaar de voorspelling is. Deze methode kan een geschikte uitkomst zijn bij transformatieproducten, waarvan verondersteld wordt dat ze in het algemeen overeenkomsten hebben met de structuur van de moederstof. Deze methode is echter gevoelig voor fouten omdat de aan- of afwezigheid van een functionele groep een groot effect kan hebben op de ionisatie-efficiëntie.²⁸

$$C_{\text{suspect}} = \frac{\text{piekoppervlakte suspect}}{\text{RF vergelijkbare stof}} \text{ (methode 1 – vergelijkbare stof op basis van structuur)}$$

3.1.2 Nabij eluerende stof

Pieke et al.³² hebben een aanpak voorgesteld op basis van de dichtstbijzijnde eluerende bekende stoffen met een bekende concentratie. De onderliggende aanname is dat stoffen die kort na elkaar elueren, ook een vergelijkbare RF hebben. Dit vereist de aanwezigheid van voldoende interne standaarden die over het hele chromatogram verspreid zijn.²⁸ Bovendien is het gradiënt en daarmee de chromatografische condities van invloed op de RF van een stof.

$$C_{\text{onbekende stof}} = \frac{\text{piekoppervlakte onbekende stof}}{\text{RF dichtstbijzijnde eluerende stof}} \text{ (methode 2 – nabij eluerende stoffen)}$$

3.1.3 Voorspellen van ionisatie-efficiëntie

Ten slotte kan de ionisatie-efficiëntie ook voorspeld worden met behulp van modellen die bijvoorbeeld met machine learning zijn ontwikkeld. Deze maken gebruik van de eigenschappen en/of structuurinformatie van de suspect- of onbekende stof om hier vervolgens de ionisatie-efficiëntie van te voorspellen. Hier zijn verschillende tools voor ontwikkeld, onder andere:

- SemiQ³³ – vrij beschikbaar via <http://trams.chem.uoa.gr/semiquantification/>,
- Quantem.co³⁴ – beschikbaar via <https://quantem.co/> en
- MS2Quant³⁵ – beschikbaar via <https://github.com/kruvelab/MS2Quant>.

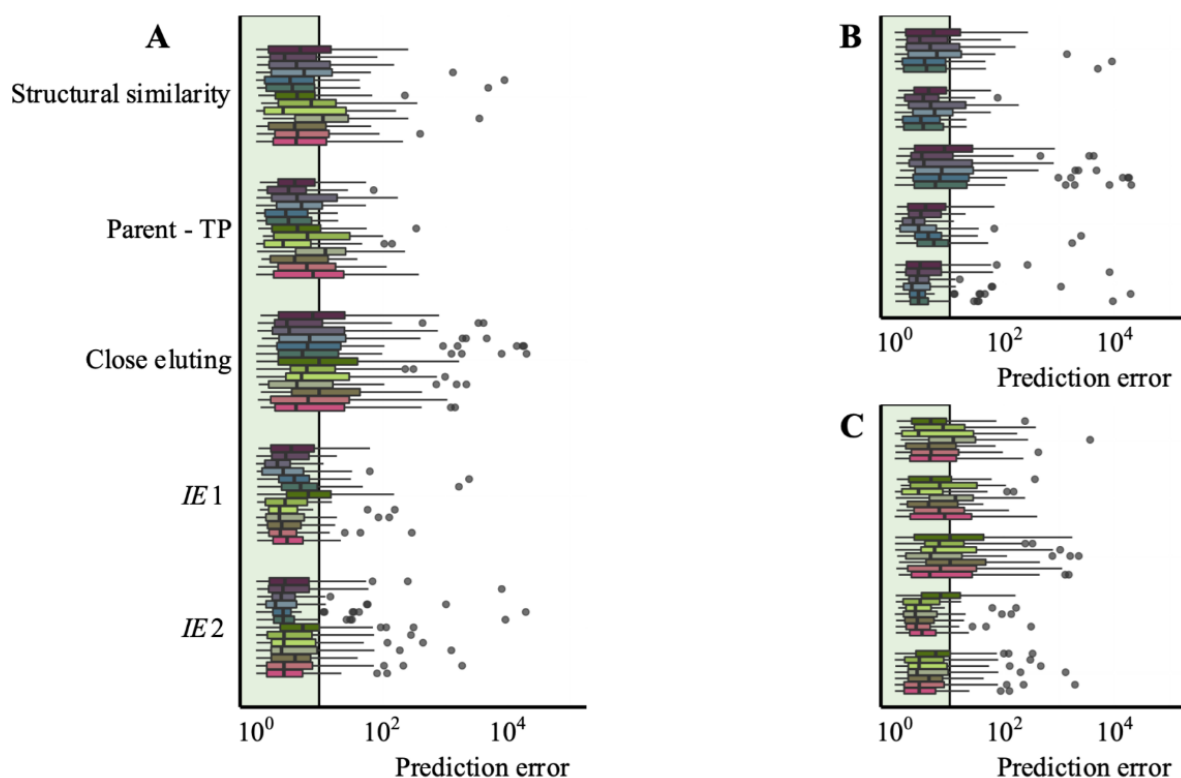
SemiQ maakt gebruik van een Quantitative Structure-Property Relationship (QSPR) aanpak om de ionisatie-efficiëntie te voorspellen. Quantem.co maakt gebruik van 'regularized random forest regressie' en MS2Quant is gebaseerd op een eXtreme Gradient Boosting (xgbTree) algoritme. De methodes zijn meetapparatuur onafhankelijk; aan de hand van een aantal bekende stoffen die met bekende concentraties gemeten zijn, wordt het model geoptimaliseerd voor de gebruikte meetapparatuur en de gebruikte meetmethode.

3.2 NORMAN ringonderzoek voor (semi-)kwantitatieve LC-HRMS non-target screening

KWR heeft deelgenomen aan het NORMAN ringonderzoek voor (semi-)kwantitatieve LC-HRMS non-target screening, georganiseerd door de universiteit van Stockholm en de universiteit van Athene.² In dit ringonderzoek werden de hierboven beschreven (semi-)kwantificeringsstrategieën getest:

1. Struct-sim: Op basis van structurele gelijkenis met bekende standaarden
2. Parent-TP: Gebruik van de moederstof om het transformatieproduct te kwantificeren
3. Close-eluting: Nabij eluerende standaarden
4. Voorspelling van ionisatie-efficiëntie
 - a. IE1: Met een online tool Quantem (<https://app.quantem.co>)³⁶
 - b. IE2: Met de online tool SemiQ v1.0.0 (<http://trams.chem.uoa.gr/semiquantification/>)³⁷

In totaal hebben 39 laboratoria resultaten ingediend. De laboratoria ontvingen elk 6 monsters, bestaande uit drie matrices (HPLC water, drinkwater en oppervlaktewater) gespiked met 45 stoffen op twee concentratieniveaus: hoog en laag. Daarnaast werden er 6 kalibratiemonsters verstrekt met 41 stoffen in HPLC water, op 6 bekende concentratieniveaus. KWR heeft deze monsters gemeten met de reguliere non-target screening LC-HRMS methode, beschreven in paragraaf 3.3.1. De definitieve resultaten zijn nog niet gepubliceerd, dit wordt verwacht in de zomer van 2024. Uit de [voorlopige resultaten](#) van het ringonderzoek blijkt dat de methodes op basis van ionisatie-efficiëntie de meest veelbelovende resultaten laten zien, d.w.z. de concentraties van 80% van de stoffen worden binnen een foutmarge van 10x voorspeld. Op basis van de resultaten en de gebruikservaring lijkt de semi-kwantificatie aanpak met Quantem de meest geschikte tool, zie Figuur 14.



Figuur 14 - Resultaten van 10 deelnemende laboratoria, gepresenteerd als boxplots voor elke semi-kwantitatieve aanpak en lab, voor één monster. A geeft de resultaten weer voor alle 10 laboratoria (in totaal 13 resultaten omdat 3 laboratoria twee resultaten hebben). B geeft de resultaten weer van laboratoria met een Orbitrap instrument, en C geeft de resultaten weer van laboratoria met een ToF instrument. Bron: (Malm & Krueve, Preliminary results update, https://kruevelab.com/wp-content/uploads/2022/06/Preliminary_results_update.pdf, Juni 2022).

² Het ringonderzoek is uitgevoerd en wordt gerapporteerd door Anneli Krueve en Nikiforos Alygizakis en gefinancierd door het NORMAN netwerk.

Quantem en de Universiteit van Athene hebben toegang gegeven tot de online semi-kwantificatie tools.

3.3 Test Quantem

Op basis van de voorlopige resultaten van het NORMAN ringonderzoek is besloten om het gebruik van Quantem te testen en vergelijken met de huidige manier van semikwantificeren: de concentraties rapporteren als interne standaard equivalenten. De interne standaarden voor de kalibratie van Quantem zijn geselecteerd op basis van advies van Jaanus Liigand (ontwikkelaar van Quantem). De teststoffen zijn gekozen op basis van de bij KWR beschikbare stoffen en mixen.

3.3.1 Methode

De stockoplossingen van de Quantem interne standaard en de te kwantificeren referentiestandaarden werden bereid in acetonitril, methanol of water, afhankelijk van de oplosbaarheid van de stoffen, met een concentratie van 1 mg/L. De stockoplossingen zijn bewaard bij -20 °C. De stockoplossing werd op twee niveaus, eindconcentraties van 0,2 (lage concentratie) en 5 µg/L (hoge concentratie), gespiked aan vier verschillende matrices: milli-Q water (MQ), drinkwater (DW), oppervlaktewater (OW) afkomstig uit het Lekkanaal en rioolwaterzuivering effluent (EFF) van rioolwaterzuivering Utrecht. De interne standaarden atrazine-d5, bentazon-d6 en benzotriazole-d4 werden aan alle monsters toegevoegd tot een eindconcentratie van 1 µg/L. De Quantem interne standaard en de overige referentiestandaarden en de concentraties staan vermeld in Bijlage I.

De LC-HRMS analyse werd in triplo uitgevoerd op een Vanquish HPLC systeem gekoppeld aan een Orbitrap Fusion Tribid massaspectrometer (Thermo Fisher Scientific). Voor de analyse werd dezelfde methode gebruikt als beschreven in Paragraaf 2.4.2. De data-analyse en integratie van de pieken werd uitgevoerd met Xcalibur™ 4.5 (Thermo Fisher Scientific). De gemiddelde piekoppervlakte van elke stof in elke matrix werd gecorrigeerd voor de oppervlakte in de blanco matrix. Vervolgens werden de piekoppervlaktes geüpload in de Quantem software. De door Quantem berekende concentraties werden geëxporteerd en verder verwerkt in R (versie 4.0.2, The R Foundation for Statistical Computing) en RStudio (versie 2023.09.1, Posit Software, PBC). Voor elke stof en spikenniveau zijn de absolute fout en de factor berekend volgens Vergelijkingen 9 en 10³⁶. Tevens werden de concentraties als interne standaard equivalenten bepaald (zie Vergelijking 11), voor de positieve ionisatiemodus werd atrazine-d5 gebruikt, voor de stoffen die gemeten zijn in de negatieve ionisatiemodus was dit bentazon-d6.

$$\text{absolute fout } (\mu\text{g/L}) = \text{voorspelde concentratie } (\mu\text{g/L}) - \text{ware concentratie } (\mu\text{g/L}) \quad (9)$$

$$\text{factor} = \max \left\{ \frac{\text{voorspelde concentratie} / \text{ware concentratie}}{\text{ware concentratie} / \text{voorspelde concentratie}} \right\} \quad (10)$$

$$\text{concentratie interne standaard equivalent } (\mu\text{g IS/L}) = \text{piekoppervlakte stof} / \frac{\text{piekoppervlakte IS}}{\text{concentratie IS}} \quad (11)$$

De absolute fout kan soms een vertekend beeld geven, daarom wordt de prioriteit gegeven aan de fout als factor. Wanneer de werkelijke concentratie van een stof bijvoorbeeld 0,1 µg/L is, en de voorspelde concentratie 1 µg/L, dan is de absolute fout slechts 0,9 µg/L, maar de voorspelling zit een factor 10 naast de werkelijke waarde. In de monitoring van de waterkwaliteit is deze ordegrrootte van belang en speelt de absolute fout een minder grote rol. De context van het getal, dus de range waarbinnen de fout ligt, is echter ook van belang.

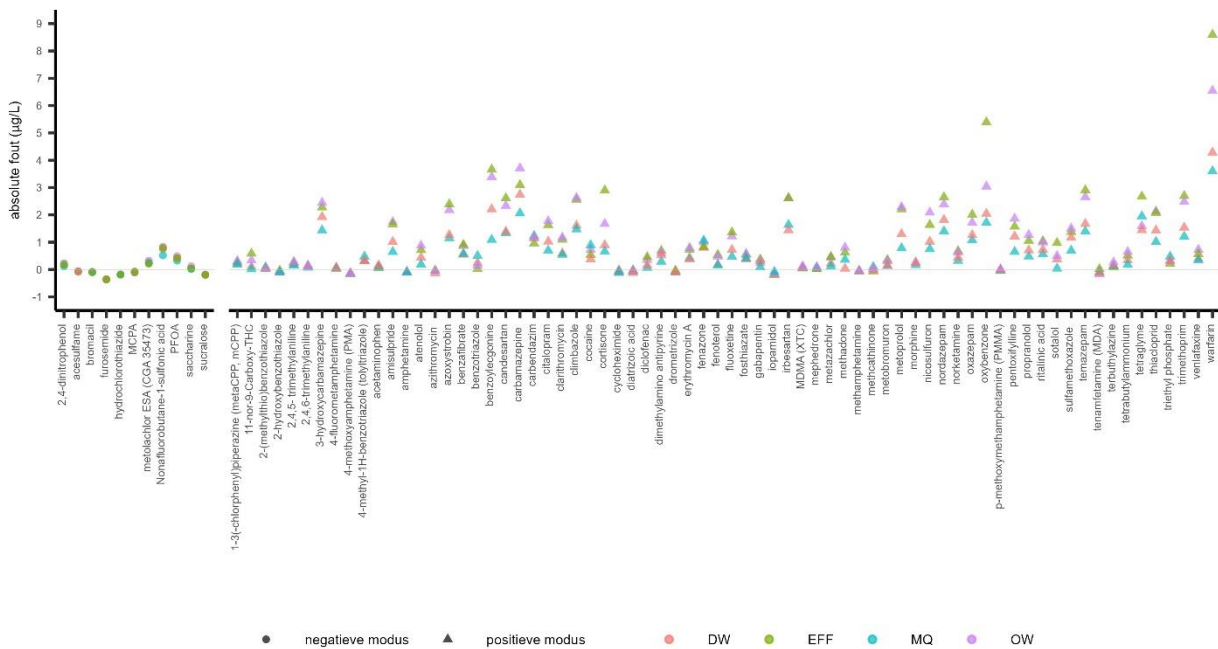
3.3.2 Resultaten

De resultaten van de Quantem test zijn weergegeven in Figuur 15 t/m Figuur 19. De datapunten van de hoge en de lage spike zijn apart weergegeven en laten duidelijk hetzelfde patroon zien. De Quantem software berekent per stof een IE factor die bepaald wordt aan de hand van het gradiënt, de mobiele fase en de piekoppervlakte van de standaarden in de matrix. In dit onderzoek is alleen met de matrix gevarieerd (milli-Q water, drinkwater, oppervlaktewater en effluent) en met het spikenniveau (laag en hoog). In bijlage II staan enkele extra figuren waarin de factor wordt uitgezet tegen de retentietijden (Figuur 21 en Figuur 22) en de ge(de)protoneerde massa (Figuur

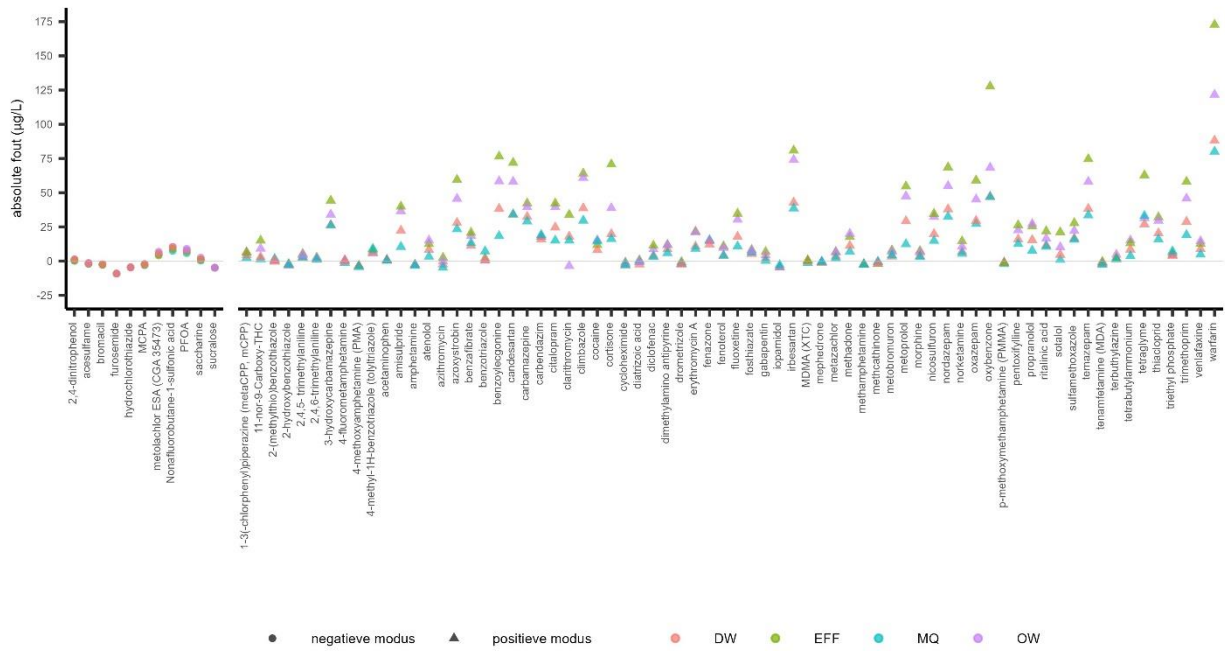
23 en Figuur 24), ook wordt de door Quantem voorspelde fout (prediction error) vergeleken met de werkelijke factor (Figuur 25 en Figuur 26).

De respons van azithromycine in milli-Q water op laag spikenniveau wijkt sterk af en wordt beschouwd als een uitbijter, dit datapunt is uit de dataset verwijderd.

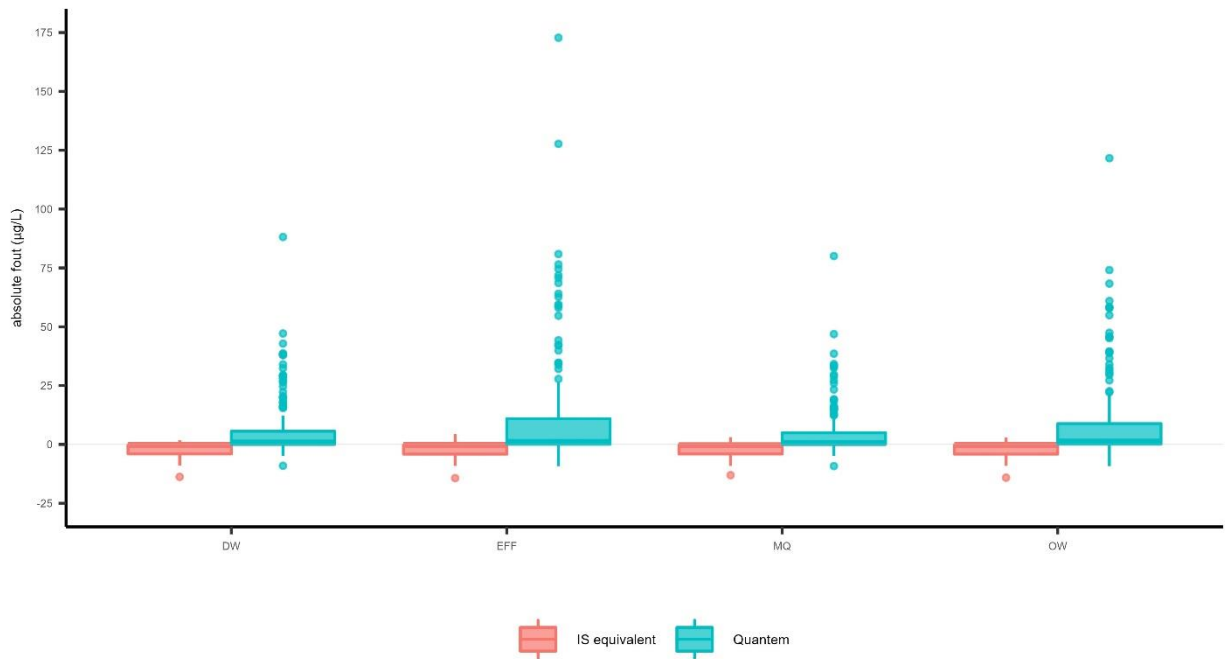
Uit Figuur 15 en Figuur 16 blijkt dat de verschillen tussen de werkelijke concentraties en de met Quantem voorspelde concentraties, de absolute fouten, uiteenlopen van klein ($\pm 0,01 \mu\text{g/L}$) tot groot ($\pm 173 \mu\text{g/L}$). Voor de interne standaard equivalenten lopen deze absolute fouten uiteen van $\pm 0,004 \mu\text{g/L}$ tot $\pm 14 \mu\text{g/L}$. De concentraties worden met behulp van Quantem vaker overschat, ($n = 499$, 77% van de geschatte concentraties zijn hoger) dan onderschat ($n = 147$, 23% van de geschatte concentraties zijn lager). In het geval van interne standaard equivalenten werden de concentraties vaker onderschat ($n = 595$, 92% van de geschatte concentraties zijn lager) dan overschat ($n = 51$, 8% van de geschatte concentraties zijn hoger). Dit is ook duidelijk zichtbaar in Figuur 17.



Figuur 15 - Verschil tussen de werkelijke concentratie in µg/L (laag spikenniveau) en de door Quantem voorspelde concentratie, per stof, matrix en ionisatiemodus.

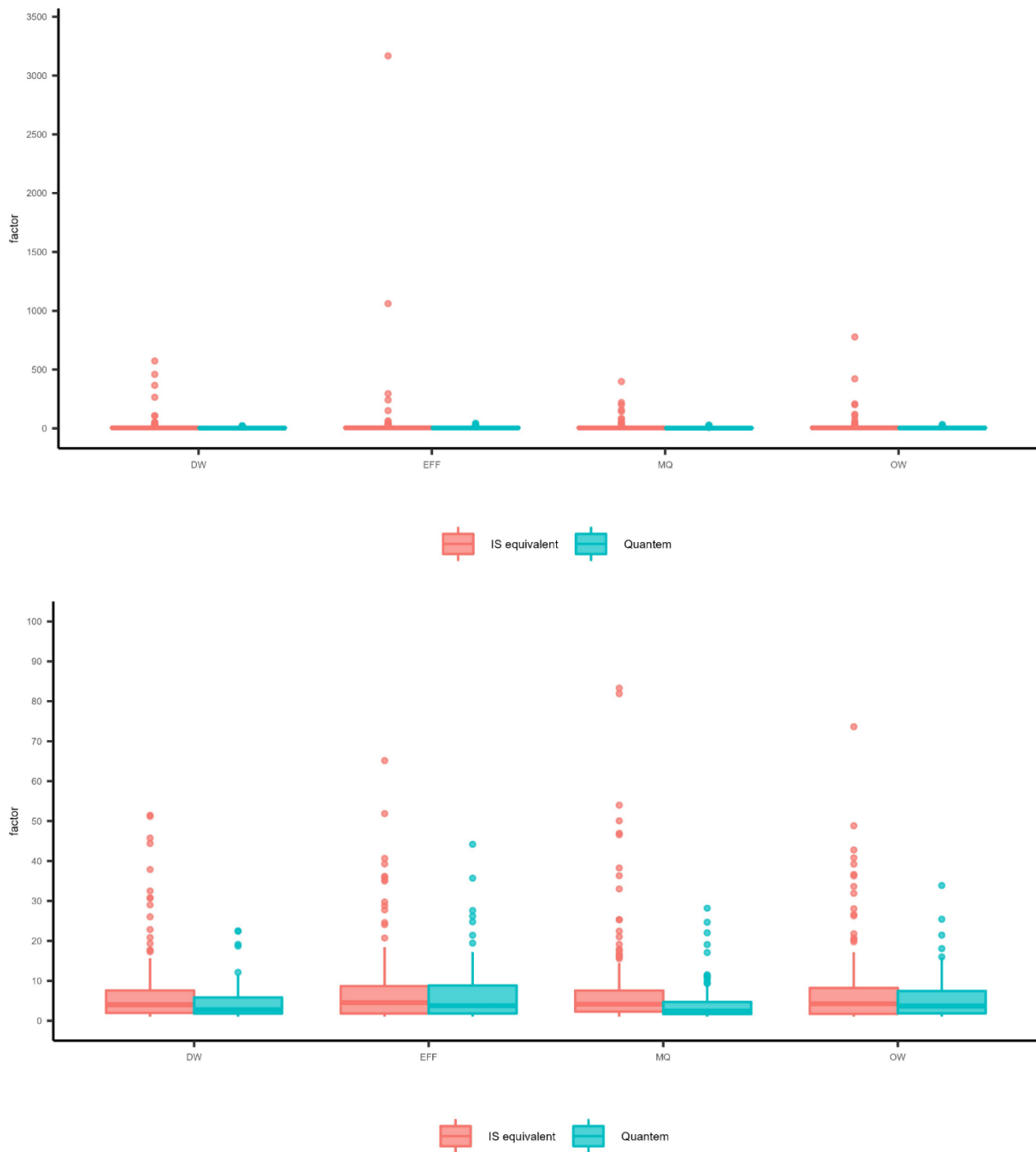


Figuur 16 - Verschil tussen de werkelijke concentratie in µg/L (hoog spikenniveau) en de door Quantem voorspelde concentratie, per stof, matrix en ionisatiemodus.



Figuur 17 – Boxplot van de absolute fout in µg/L van de Quantem resultaten vergeleken met absolute fout van de berekende interne standaard equivalenten. Voor alle matrices is het verschil significant (Wilcoxon-Signed rank test, p-waarde < 0,05).

Figuur 18 en Figuur 19 laten de verschillen tussen de voorspelling van Quantem en de werkelijke concentratie op beide spikenniveaus zien, deze afwijkingen zijn gepresenteerd als factor van de werkelijke concentratie (zie Vergelijking 10). In deze figuren is te zien dat deze wederom ver uiteenlopen van klein ($\pm 1x$) tot groot ($\pm 44x$). Dit geldt zowel voor het hoge spikenniveau als het lage spikenniveau. Voor de interne standaard equivalenten ligt dit veel



Figuur 20 – Boxplot van de factor van de Quantem resultaten vergeleken met de factor van de berekende interne standaard equivalenten (boven). Zoom in op de boxplots (onder). Voor MQ en DW is het verschil tussen de verschillende voorspellingen significant (Wilcoxon Signed rank test, p-waarde < 0,05).

3.3.3 Discussie

Uit de resultaten blijkt dat de concentratievoorspellingen niet altijd even nauwkeurig zijn. In het geval van voorspellingen met Quantem worden de concentraties vaker overschat, waar in het geval van interne standaard equivalenten de concentraties vaker onderschat worden. In dit onderzoek zijn alleen atrazine-d5 en bentazon-d6 gebruikt om de interne standaard equivalenten te bepalen. Wanneer andere interne standaarden gebruikt worden kunnen de resultaten anders zijn, deze resultaten zijn dus niet extrapoleerbaar naar alle schattingen op basis van interne standaarden. De resultaten van de negatieve ionisatiemodus kunnen niet gelijkwaardig vergeleken worden met de positieve ionisatiemodus omdat het aantal gemeten stoffen verschilt.

Vanuit het drinkwaterperspectief brengt het onderschatten van de concentratie een groter toxicologisch risico met zich mee dan het overschatten van de concentratie. Atrazine-d5 is een stof met een goede IE³⁶ en bentazon-d6 is bij KWR een van de gevoeligste componenten die in de negatieve modus gemeten wordt, waardoor de kans groter wordt dat andere stoffen een minder goede IE hebben en er dus vaker een onderschatting van de concentratie plaatsvindt. Wanneer gekeken wordt naar het verschil als factor van de werkelijke concentratie, dan valt op dat er meer extreme uitbijters in de interne standaard equivalenten voorspelling zit. Wanneer een factor 2 wordt aangehouden als een acceptabele afwijking dan valt 32% van de Quantem voorspellingen daarbinnen, voor de interne standaard equivalenten is dit 25%. Wanneer uitgegaan wordt van een factor 5 dan valt 68% van de Quantem voorspellingen daarbinnen, voor de interne standaard equivalenten is dit 54%. Op basis van de resultaten van dit onderzoek lijken de voorspellingen van Quantem nauwkeuriger dan de schatting van de concentratie met behulp van de interne standaarden atrazine-d5 (in positieve modus) en bentazon-d6 (in negatieve modus). Er zijn geen duidelijke trends zichtbaar wanneer de afwijkingen uitgezet worden tegenover de retentietijd en de m/z waarde van het moleculair ion ([M+H]⁺ of [M-H]⁻).

Quantem is een commerciële tool, welke zeer gebruiksvriendelijk is maar waar wel kosten aan verbonden zijn, in tegenstelling tot het berekenen van interne standaard equivalenten. De kosten worden berekend per voorspelde concentratie (± €1 per analiet per meting), er worden verschillende pakketten aangeboden van verschillende grootte. De prijs per analiet gaat omlaag naarmate er meer analieten voorspeld worden. De input van Quantem bestaat uit een CSV of Excel bestand met daarin de piekoppervlakte, retentietijd en SMILES code per stof. De output bestaat uit een CSV bestand met dezelfde gegevens en een voorspelde concentratie daaraan toegevoegd. Dit is eenvoudig te verwerken in vervolgstappen.

Ten slotte is in dit onderzoek slechts een beperkte set aan organische microverontreinigingen getest in slechts twee concentratieniveaus (0,2 µg/L en 5 µg/L), dit kan een wat vertekend beeld van de modelprestaties geven. De schattingen op basis van de interne standaarden zouden kunnen verbeteren wanneer er meerdere interne standaarden gebruikt worden met verschillende eigenschappen en retentietijden.

3.4 Conclusie semi-kwantitatieve modellen

In dit BTO onderzoek zijn een aantal modellen getest in het NORMAN ringonderzoek voor semi-kwantitatieve NTS. Naar aanleiding van de voorlopige resultaten van dit ringonderzoek is de software Quantem nader getest voor toepassing in waterkwaliteitsmonitoring en vergeleken met de huidige aanpak: concentraties schatten als interne standaard equivalenten. Op basis van de beschreven resultaten lijkt Quantem nauwkeurigere resultaten (32% van de voorspellingen valt binnen een factor 2, en 68% valt binnen een factor 5) te genereren dan de interne standaard equivalenten methode (25% van de voorspellingen valt binnen een factor 2, en 54% valt binnen een factor 5). Het gebruik van Quantem brengt kosten met zich mee, daarom is het advies om het toe te passen bij projecten/vraagstukken waar de semikwantificatie echt van toegevoegde waarde is. Quantem zou ook gebruikt kunnen worden als prioriteringstool. De semi-kwantitatieve modellen voor non-target en suspect screening zijn volop in ontwikkeling en de verwachting is dat deze in de komende jaren sterk zullen verbeteren. Uiteraard zal er altijd een bepaalde mate van onzekerheid gepaard gaan met voorspellingen, gezien de grote verscheidenheid van factoren die invloed hebben op de IE van een stof. Ten slotte kunnen de verder en nieuw ontwikkelde tools nader onderzocht worden zodra de resultaten van het ringonderzoek bekend zijn en de aanbevelingen daaruit ook in ogenschouw genomen kunnen worden.

4 Referenties

- (1) Hollender, J.; Schymanski, E. L.; Singer, H. P.; Ferguson, P. L. Nontarget Screening with High Resolution Mass Spectrometry in the Environment: Ready to Go? *Environ Sci Technol* **2017**, *51* (20), 11505-11512. DOI: 10.1021/acs.est.7b02184.
- (2) Rigano, F.; Arigo, A.; Oteri, M.; La Tella, R.; Dugo, P.; Mondello, L. The retention index approach in liquid chromatography: An historical review and recent advances. *J Chromatogr A* **2021**, *1640*, 461963. DOI: 10.1016/j.chroma.2021.461963.
- (3) Staatssecretaris van Infrastructuur en Milieu. Drinkwaterregeling. In *BJZ2011046947*, Overheid.nl, 2011.
- (4) Programmteam Water. Protocol voor monitoring en toetsing drinkwaterbronnen KRW. Informatiepunt Leefomgeving, Ed.; 2015.
- (5) Oss, M.; Kruve, A.; Herodes, K.; Leito, I. Electrospray Ionization Efficiency Scale of Organic Compounds. *Analytical Chemistry* **2010**, *82* (7), 2865-2872. DOI: 10.1021/ac902856t.
- (6) Banerjee, S.; Mazumdar, S. Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. *Int J Anal Chem* **2012**, *2012*, 282574. DOI: 10.1155/2012/282574.
- (7) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J. Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ Sci Technol* **2014**, *48* (4), 2097-2098. DOI: 10.1021/es5002105.
- (8) Bonini, P.; Kind, T.; Tsugawa, H.; Barupal, D. K.; Fiehn, O. Retip: Retention Time Prediction for Compound Annotation in Untargeted Metabolomics. *Anal Chem* **2020**, *92* (11), 7515-7522. DOI: 10.1021/acs.analchem.9b05765.
- (9) Naylor, B. C.; Catrow, J. L.; Maschek, J. A.; Cox, J. E. QSRR Automator: A Tool for Automating Retention Time Prediction in Lipidomics and Metabolomics. *Metabolites* **2020**, *10* (6). DOI: 10.3390/metabo10060237.
- (10) Aalizadeh, R.; Thomaidis, N. S.; Bletsou, A. A.; Gago-Ferrero, P. Quantitative Structure-Retention Relationship Models To Support Nontarget High-Resolution Mass Spectrometric Screening of Emerging Contaminants in Environmental Samples. *J Chem Inf Model* **2016**, *56* (7), 1384-1398. DOI: 10.1021/acs.jcim.5b00752.
- (11) McEachran, A. D.; Mansouri, K.; Newton, S. R.; Beverly, B. E. J.; Sobus, J. R.; Williams, A. J. A comparison of three liquid chromatography (LC) retention time prediction models. *Talanta* **2018**, *182*, 371-379. DOI: 10.1016/j.talanta.2018.01.022.
- (12) *Estimation Programs Interface Suite™ for Microsoft® Windows*; United States Environmental Protection Agency: Washington, DC, USA, 2020. (accessed).
- (13) Mansouri, K.; Grulke, C. M.; Judson, R. S.; Williams, A. J. OPERA models for predicting physicochemical properties and environmental fate endpoints. *J Cheminform* **2018**, *10* (1), 10. DOI: 10.1186/s13321-018-0263-1.
- (14) Yap, C. W. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. *J Comput Chem* **2011**, *32* (7), 1466-1474. DOI: 10.1002/jcc.21707.
- (15) Golbraikh, A.; Tropsha, A. Beware of q²! *Journal of Molecular Graphics and Modelling* **2002**, *2002* (20), 269-276. DOI: [https://doi.org/10.1016/S1093-3263\(01\)00123-1](https://doi.org/10.1016/S1093-3263(01)00123-1).
- (16) Ruttkies, C.; Schymanski, E. L.; Wolf, S.; Hollender, J.; Neumann, S. MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *J Cheminform* **2016**, *8*, 3. DOI: 10.1186/s13321-016-0115-9.
- (17) Aicheler, F.; Li, J.; Hoene, M.; Lehmann, R.; Xu, G.; Kohlbacher, O. Retention Time Prediction Improves Identification in Nontargeted Lipidomics Approaches. *Analytical Chemistry* **2015**, *87* (15), 7698-7704. DOI: <https://doi.org/10.1021/acs.analchem.5b01139>.
- (18) Helmus, R.; ter Laak, T. L.; van Wezel, A. P.; de Voogt, P.; Schymanski, E. L. patRoom: open source platform for environmental mass spectrometry based non-target screening. *Journal of Cheminformatics* **2021**. DOI: <https://doi.org/10.1186/s13321-020-00477-w>.
- (19) Witting, M.; Bocker, S. Current status of retention time prediction in metabolite identification. *J Sep Sci* **2020**, *43* (9-10), 1746-1754. DOI: 10.1002/jssc.202000060.
- (20) Ehrmann, B. M.; Henriksen, T.; Cech, N. B. Relative importance of basicity in the gas phase and in solution for determining selectivity in electrospray ionization mass spectrometry. *J Am Soc Mass Spectrom* **2008**, *19* (5), 719-728. DOI: 10.1016/j.jasms.2008.01.003.
- (21) Leito, I.; Herodes, K.; Huopola, M.; Virro, K.; Künnapas, A.; Kruve, A.; Tanner, R. Towards the electrospray ionization mass spectrometry ionization efficiency scale of organic compounds. *Rapid Communications in Mass Spectrometry* **2008**, *22* (3), 379-384. DOI: <https://doi.org/10.1002/rcm.3371>.

- (22) Krueve, A.; Kaupmees, K.; Liigand, J.; Leito, I. Negative electrospray ionization via deprotonation: predicting the ionization efficiency. *Analytical Chemistry* **2014**, *86* (10), 4822-4830. DOI: 10.1021/ac404066v From NLM PubMed-not-MEDLINE.
- (23) Krueve, A. Strategies for Drawing Quantitative Conclusions from Nontargeted Liquid Chromatography-High-Resolution Mass Spectrometry Analysis. *Anal Chem* **2020**, *92* (7), 4691-4699. DOI: 10.1021/acs.analchem.9b03481 From NLM PubMed-not-MEDLINE.
- (24) Chalcraft, K. R.; Lee, R.; Mills, C.; Britz-McKibbin, P. Virtual Quantification of Metabolites by Capillary Electrophoresis-Electrospray Ionization-Mass Spectrometry: Predicting Ionization Efficiency Without Chemical Standards. *Analytical Chemistry* **2009**, *81*, 2506-2515. DOI: <https://doi.org/10.1021/ac802272u>.
- (25) Liigand, J.; Krueve, A.; Leito, I.; Girod, M.; Antoine, R. Effect of mobile phase on electrospray ionization efficiency. *J Am Soc Mass Spectrom* **2014**, *25* (11), 1853-1861. DOI: 10.1007/s13361-014-0969-x.
- (26) Liigand, J.; Laaniste, A.; Krueve, A. pH Effects on Electrospray Ionization Efficiency. *J Am Soc Mass Spectrom* **2017**, *28* (3), 461-469. DOI: 10.1007/s13361-016-1563-1.
- (27) Ojakivi, M.; Liigand, J.; Krueve, A. Modifying the Acidity of Charged Droplets. *ChemistrySelect* **2018**, *3* (1), 335-338. DOI: 10.1002/slct.201702269 (accessed 2019/08/27).
- (28) Malm, L.; Palm, E.; Souihi, A.; Plassmann, M.; Liigand, J.; Krueve, A. Guide to Semi-Quantitative Non-Targeted Screening Using LC/ESI/HRMS. *Molecules* **2021**, *26* (12). DOI: 10.3390/molecules26123524.
- (29) Liigand, J.; Krueve, A.; Leito, I.; Girod, M.; Antoine, R. Effect of mobile phase on electrospray ionization efficiency. *Journal of The American Society for Mass Spectrometry* **2014**, *25* (11), 1853-1861. DOI: 10.1007/s13361-014-0969-x.
- (30) Liigand, J.; Krueve, A.; Liigand, P.; Laaniste, A.; Girod, M.; Antoine, R.; Leito, I. Transferability of the Electrospray Ionization Efficiency Scale between Different Instruments. *Journal of The American Society for Mass Spectrometry* **2015**, *26* (11), 1923-1930. DOI: <https://doi.org/10.1021/jasms.8b04919>.
- (31) Sepman, H.; Malm, L.; Peets, P.; Krueve, A. Scientometric review: Concentration and toxicity assessment in environmental non-targeted LC/HRMS analysis. *Trends in Environmental Analytical Chemistry* **2023**, *40*. DOI: 10.1016/j.teac.2023.e00217.
- (32) Pieke, E. N.; Granby, K.; Trier, X.; Smedsgaard, J. A framework to estimate concentrations of potentially unknown substances by semi-quantification in liquid chromatography electrospray ionization mass spectrometry. *Anal Chim Acta* **2017**, *975*, 30-41. DOI: 10.1016/j.aca.2017.03.054.
- (33) Aalizadeh, R.; Nikolopoulou, V.; Alygizakis, N.; Slobodnik, J.; Thomaidis, N. S. A novel workflow for semi-quantification of emerging contaminants in environmental samples analyzed by LC-HRMS. *Analytical and Bioanalytical Chemistry* **2022**. DOI: <https://doi.org/10.1007/s00216-022-04084-6>.
- (34) Liigand, J.; Wang, T.; Kellogg, J.; Smedsgaard, J.; Cech, N.; Krueve, A. Quantification for non-targeted LC/MS screening without standard substances. *Sci Rep* **2020**, *10* (1), 5808. DOI: 10.1038/s41598-020-62573-z.
- (35) Sepman, H.; Malm, L.; Peets, P.; MacLeod, M.; Martin, J.; Breitholtz, M.; Krueve, A. Bypassing the Identification: MS2Quant for Concentration Estimations of Chemicals Detected with Nontarget LC-HRMS from MS(2) Data. *Anal Chem* **2023**, *95* (33), 12329-12338. DOI: 10.1021/acs.analchem.3c01744.
- (36) Liigand, J.; Wang, T.; Kellogg, J.; Smedsgaard, J.; Cech, N.; Krueve, A. Quantification for non-targeted LC/MS screening without standard substances. *Scientific Reports Nature Research* **2020**, *10* (1), 5808. DOI: 10.1038/s41598-020-62573-z From NLM PubMed-not-MEDLINE.
- (37) Aalizadeh, R.; Panara, A.; Thomaidis, N. S. Development and Application of a Novel Semi-quantification Approach in LC-QToF-MS Analysis of Natural Products. *Journal of The American Society for Mass Spectrometry* **2021**, *32* (6), 1412-1423. DOI: <https://doi.org/10.1021/jasms.1c00032>.

I Overzicht stoffen

Tabel 11 – Referentiestandaarden en concentraties, positieve ionisatiemodus.

Naam	Brutoformule	[M+H] ⁺ (m/z)	RT (min)	Spike laag (µg/L)	Spike hoog (µg/L)
iopamidol	C17H22I3N3O8	777,86137	2,26	0,209	5,234
morphine	C17H19NO3	286,14377	2,45	0,206	5,162
diatrizoic acid	C11H9I3N2O4	614,77694	2,48	0,193	4,824
atenolol	C14H22N2O3	267,17032	3,55	0,194	4,850
sotalol	C12H20N2O3S	273,12674	4,35	0,200	5,008
acetaminophen	C8H9NO2	152,07061	4,55	0,201	5,015
dimethylamino antipyrine	C13H17N3O	232,14444	5,52	0,196	4,891
methcathinone	C10H13NO	164,10699	5,90	0,202	5,042
carbendazim	C9H9N3O2	192,07675	6,25	0,199	4,966
gabapentin	C9H17NO2	172,13320	6,35	0,405	10,117
fenoterol	C17H21NO4	304,15433	6,60	0,195	4,866
amphetamine	C9H13N	136,11208	6,86	0,200	5,007
tenamfetamine (MDA)	C10H13NO2	180,10191	7,24	0,199	4,985
methamphetamine	C10H15N	150,12773	7,42	0,200	5,012
4-methoxyamphetamine (PMA)	C10H15NO	166,12264	7,50	0,203	5,066
MDMA (XTC)	C11H15NO2	194,11756	7,66	0,202	5,057
trimethoprim	C14H18N4O3	291,14517	7,72	0,201	5,035
tetraglyme	C10H22O5	223,15400	7,76	0,200	4,989
benzotriazole	C6H5N3	120,05562	7,94	0,611	15,284
p-methoxymethamphetamine (PMMA)	C11H17NO	180,13829	7,94	0,196	4,906
4-fluorometamphetamine	C10H14FN	168,11830	8,11	0,197	4,923
amisulpride	C17H27N3O4S	370,17950	8,13	0,204	5,092
norketamine	C12H14ClNO	224,08367	8,16	0,203	5,073
mephedrone	C11H15NO	178,12264	8,19	0,197	4,929
ritalinic acid	C13H17NO2	220,13321	8,20	0,201	5,027
benzoylecgonine	C16H19NO4	290,13868	8,36	0,198	4,960
2,4,5- trimethylaniline	C9H13N	136,11208	8,53	0,199	4,966
fenazone	C11H12N2O	189,10224	8,62	0,210	5,239
metoprolol	C15H25NO3	268,19072	9,29	0,400	10,003
2,4,6-trimethylaniline	C9H13N	136,11208	9,37	0,194	4,850
pentoxifylline	C13H18N4O3	279,14517	9,43	0,201	5,031
1-3-(chlorphenyl)piperazine (metaCPP, mCPP)	C10H13ClN2	197,08400	9,76	0,200	4,994
azithromycin	C38H72N2O12	749,51580	9,93	0,199	4,979
4-methyl-1H-benzotriazole (tolyltriazole)	C7H7N3	134,07127	9,97	0,202	5,047
cocaine	C17H21NO4	304,15433	10,25	0,194	4,857
sulfamethoxazole	C10H11N3O3S	254,05939	10,71	0,200	5,012
venlafaxine	C17H27NO2	278,21146	10,83	0,202	5,049
triethyl phosphate	C6H15O4P	183,07807	10,94	0,204	5,100
cycloheximide	C15H23NO4	282,16998	11,58	0,198	4,947
2-hydroxybenzothiazole	C7H5NOS	152,01646	11,59	0,200	5,012
3-hydroxycarbamazepine	C15H12N2O2	253,09715	11,64	0,211	5,271
propranolol	C16H21NO2	260,16451	11,78	0,379	9,468
thiacloprid	C10H9ClN4S	253,03092	11,85	0,394	9,858
erythromycin A	C37H67NO13	734,46852	12,18	0,194	4,847
nicosulfuron	C15H18N6O6S	411,10813	12,22	0,200	5,006
cortisone	C21H28O5	361,20095	12,27	0,203	5,079
citalopram	C20H21FN2O	325,17107	12,65	0,196	4,890

carbamazepine	C15H12N2O	237,10224	13,22	0,399	9,974
oxazepam	C15H11ClN2O2	287,05818	13,88	0,190	4,745
clarithromycin	C38H69N1O13	748,48417	13,90	0,199	4,970
irbesartan	C25H28N6O	429,23974	14,07	0,199	4,987
climbazole	C15H17ClN2O2	293,10513	14,11	0,199	4,976
methadone	C21H27NO	310,21654	14,31	0,197	4,932
candesartan	C24H20N6O3	441,16696	14,31	0,193	4,823
nordazepam	C15H11ClN2O	271,06327	14,44	0,197	4,913
fluoxetine	C17H18F3NO	310,14133	14,51	0,200	5,012
fosthiazate	C9H18NO3PS2	284,05385	14,61	0,204	5,108
tetrabutylammonium	C16H36N+	242,28423	14,99	0,214	5,351
temazepam	C16H13ClN2O2	301,07383	15,20	0,198	4,943
metobromuron	C9H11BrN2O2	259,00767	15,56	0,202	5,049
metazachlor	C14H16ClN3O	278,10547	15,85	0,200	4,992
benzafibrate	C19H20ClNO4	362,11536	15,90	0,194	4,857
terbutylazine	C9H16ClN5	230,11670	16,81	0,201	5,016
warfarin	C19H16O4	309,11214	16,81	0,199	4,978
2-(methylthio)benzothiazole	C8H7NS2	182,00927	17,36	0,195	4,880
azoxystrobin	C22H17N3O5	404,12410	17,64	0,199	4,986
diclofenac	C14H11Cl2NO2	296,02396	18,31	0,199	4,979
oxybenzone	C14H12O3	229,08592	19,30	0,203	5,069
11-nor-9-Carboxy-THC	C21H28O4	345,20604	21,32	0,200	5,007
drometrizole	C13H11N3O	226,09749	21,98	0,204	5,104

Tabel 12 – Referentiestandaarden en concentraties, negatieve ionisatiemodus.

Naam	Brutoformule	[M-H] ⁻ (m/z)	RT (min)	Spike laag (µg/L)	Spike hoog (µg/L)
acesulfame	C4H5NO4S	161,98665	3,29	0,165	4,137
saccharine	C7H5NO3S	181,99174	5,34	0,202	5,040
hydrochlorothiazide	C7H8ClN3O4S2	295,95719	7,22	0,200	5,012
sucralose	C12H19Cl3O8	395,00727	8,16	0,204	5,098
metolachlor ESA (CGA 35473)	C15H23NO5S	328,12242	11,37	0,200	5,011
bromacil	C9H13BrN2O2	259,00876	12,41	0,198	4,955
2,4-dinitrophenol	C6H4N2O5	183,00474	13,17	0,200	5,008
furosemide	C12H11ClN2O5S	329,00044	13,41	0,404	10,102
nonafluorobutane-1-sulfonic acid	C4HF9O3S	298,94299	13,72	0,212	5,290
MCPA	C9H9ClO3	199,01675	15,29	0,202	5,054
PFOA	C8HF15O2	412,96643	16,77	0,187	4,686

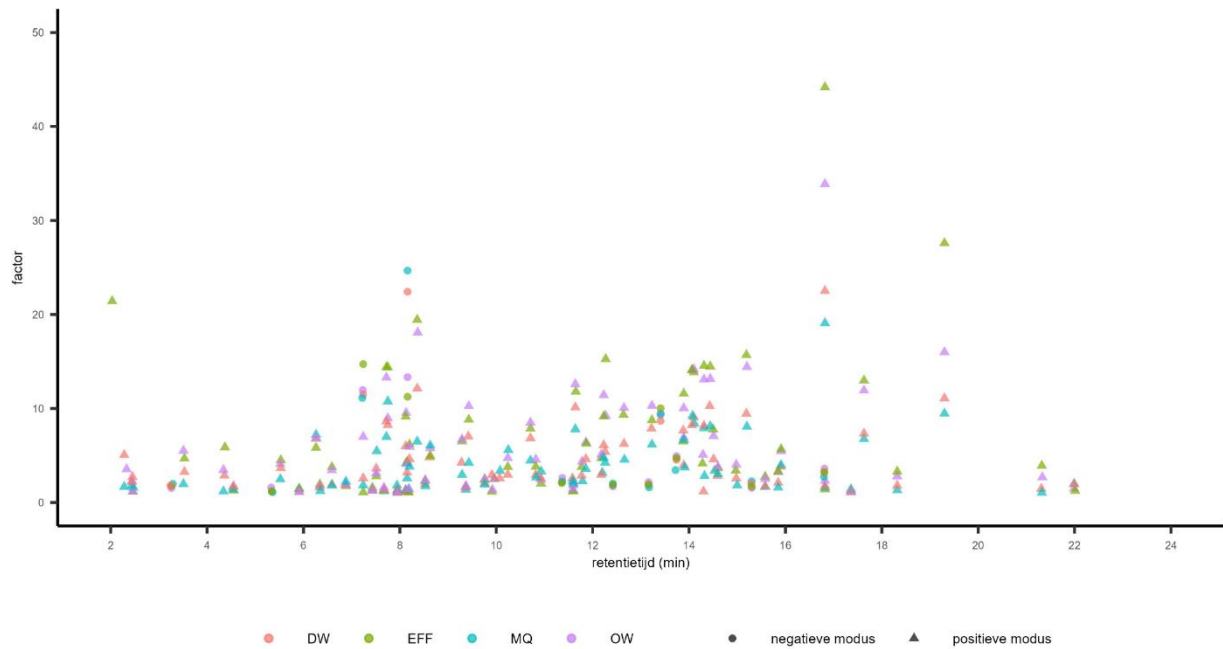
Tabel 13 – Standaarden voor Quantem, positieve ionisatiemodus

Naam	Brutoformule	[M+H] ⁺ (m/z)	RT (min)	Concentratie (µg/L)
tetrahexylammonium	C24H52N+	354,40943	21,94	2,226
tripropylamine	C9H21N	144,17468	5,84	2,014
tetraethylammonium	C8H20N+	130,15903	2,27	2,371
4-methoxypyridine	C6H7NO	110,06004	2,27	2,488
indazole	C7H6N2	119,06037	10,78	2,193
benzamide	C7H7NO	122,06004	6,92	2,219
caffeine	C8H10N4O2	195,08765	6,82	2,307
atrazine-d5	C8H9D5ClN5	221,13243	14,43	1,030
benzotriazool-d4	C6HD4N3	124,08073	7,86	1,075

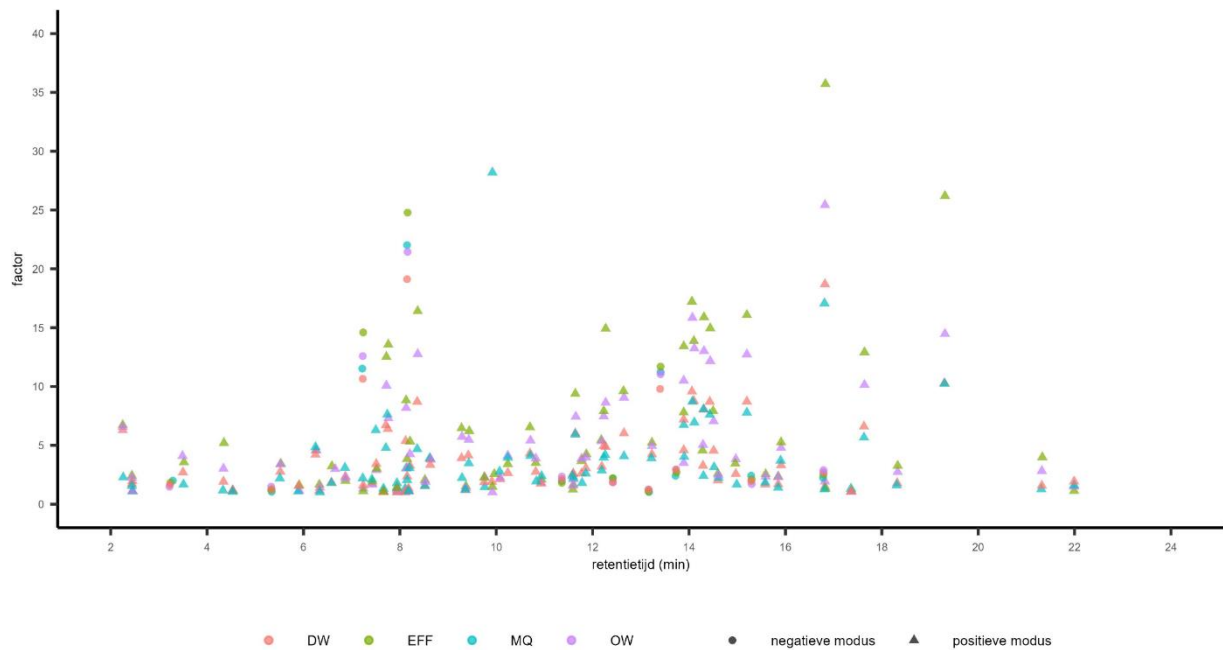
Tabel 14 – Standaarden voor Quantem, negatieve ionisatiemodus

Naam	Brutoformule	[M-H] ⁻ (m/z)	RT (min)	Concentratie (µg/L)
salicylic acid	C7H6O3	137,02442	11,42	2,411
4-nitrophenol	C6H5NO3	138,01967	11,64	2,705
3,5-dinitrosalicylic acid	C7H4N2O7	226,99457	12,07	2,218
perfluorooctanesulfonic acid	C8HF17O3S	498,93022	19,64	2,418
bentazon-d6	C10H6D6N2O3S	245,08724	14,40	1,000
benzotriazool-d4	C6HD4N3	122,06618	7,86	1,075

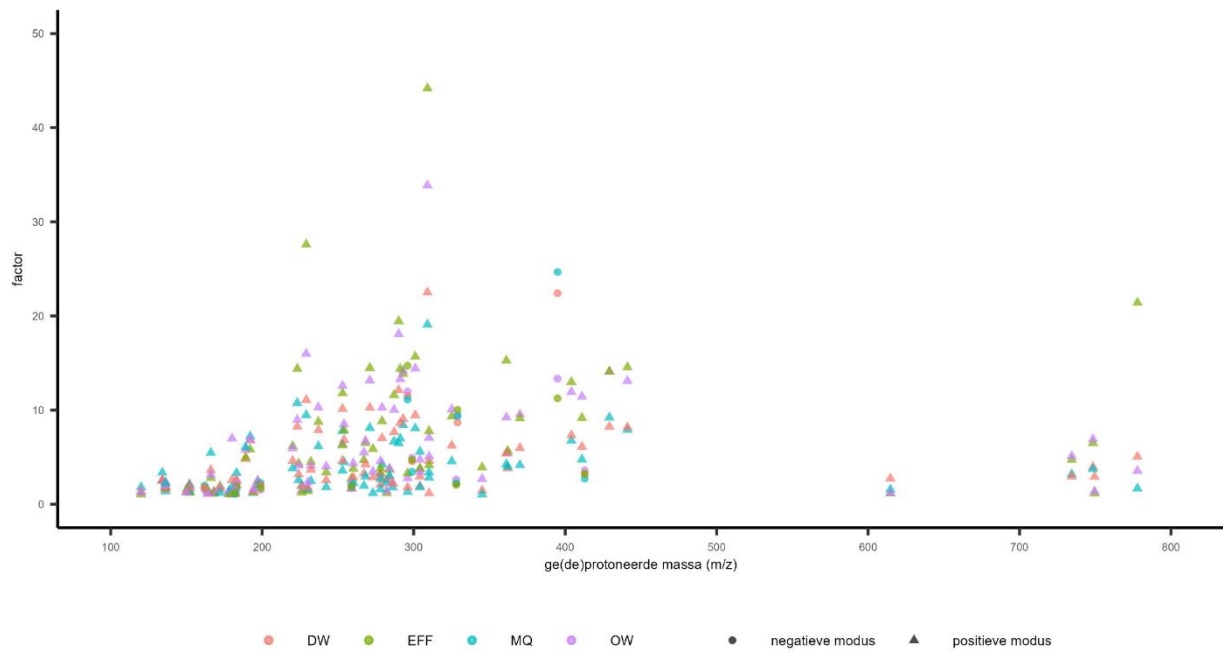
II Resultaten Quantem



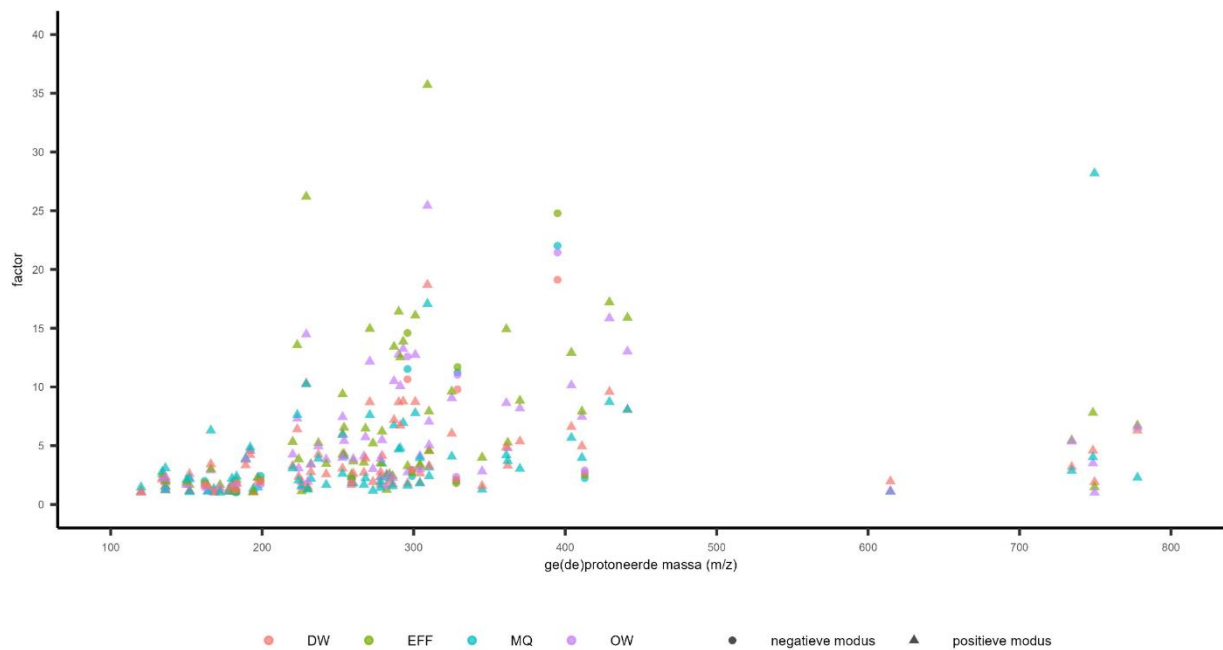
Figuur 21 – Verschil tussen de voorspelde en de ware concentratie als factor uitgezet tegen de retentietijden, **laag spikenniveau**.



Figuur 22 – Verschil tussen de voorspelde en de ware concentratie als factor uitgezet tegen de retentietijden, **hoog spikenniveau**.



Figuur 23 – Verschil tussen de voorspelde en de ware concentratie als factor uitgezet tegen de $[M+H]^+$ massa (in het geval van positieve ionisatie) en $[M-H]^-$ massa (in het geval van negatieve ionisatie), **laag spikenniveau**.



Figuur 24 – Verschil tussen de voorspelde en de ware concentratie als factor uitgezet tegen de $[M+H]^+$ massa (in het geval van positieve ionisatie) en $[M-H]^-$ massa (in het geval van negatieve ionisatie), **hoog spikenniveau**.

