

BTO 2016.061 | Juli 2016

BTO rapport

Relatie tussen ontwerp
en bedrijfsvoering van
grondwaterzuivering en
biologische activiteit

BTO

Relatie tussen ontwerp en bedrijfsvoering van grondwaterzuivering en biologische activiteit

BTO 2016.061 | Juli 2016

Opdrachtnummer

400554-091

Projectmanager

M. Hootsmans; L. Hornstra

Opdrachtgever

BTO - Thematisch onderzoek - Biologische activiteit

Kwaliteitsborger(s)

P. van der Wielen

Auteur(s)

L. Palmen, E. Vonk, D. Vries

Verzonden aan

Themagroep Biologische Activiteit

Jaar van publicatie
2016

Meer informatie

ir. L.J. Palmen MTD
T 756
E luc.palmen@kwrwater.nl

Keywords

Grondwater zuivering, biologische
stabiliteit, nagroei

PO Box 1072
3430 BB Nieuwegein
The Netherlands

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

KWR Watercycle
Research
Institute

BTO 2016.061 | December 2015 © KWR

Alle rechten voorbehouden.

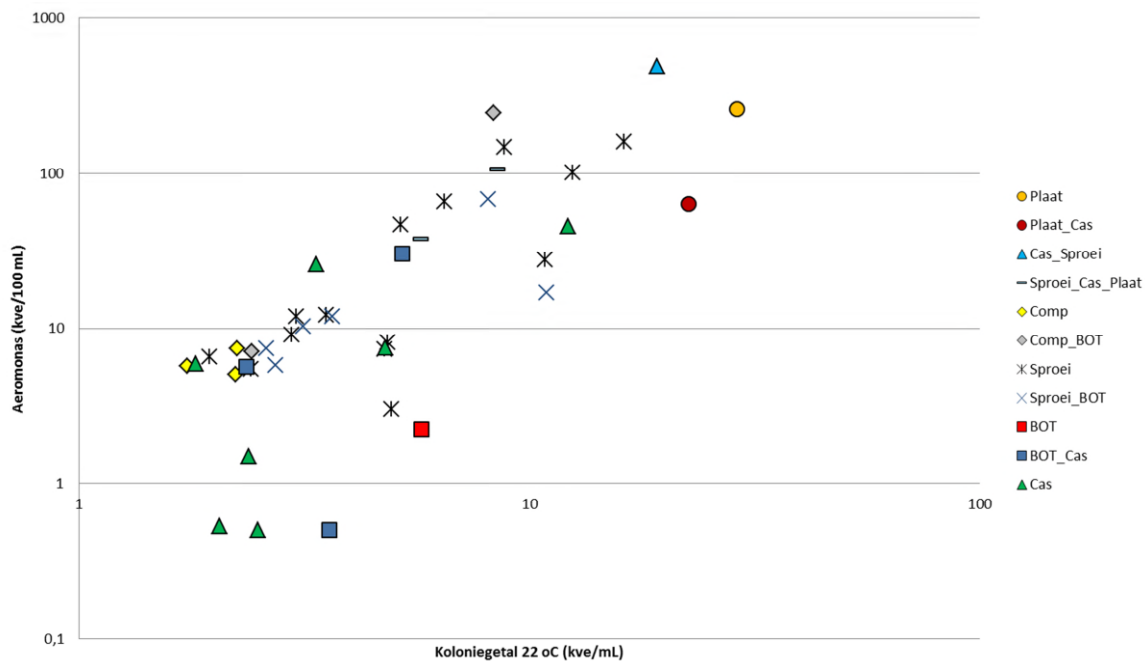
Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of enig andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

BTO Managementsamenvatting

Relaties tussen ontwerp en bedrijfsvoering van grondwaterzuivering en biologische activiteit van drinkwater

Auteur(s) ir. L.J. (Luc) Palmen, MTD, E. (Erwin) Vonk, MSc, dr. ir. D. (Dirk) Vries

De productie van biologisch stabiel water is van belang om nagroei in het distributienet te voorkomen. Naast de ruwwaterkwaliteit is ook het ontwerp en de bedrijfsvoering van de zuivering van grondwaterproductielocaties van invloed op de biologische stabiliteit. In dit onderzoek zijn de waterkwaliteitsgegevens en de ontwerp- en bedrijfsvoeringsgerelateerde eigenschappen van de beluchting en snelfiltratie van 43 grondwater productielocaties in een database verwerkt. Met verschillende analysemethoden is gezocht naar mogelijke relaties tussen deze eigenschappen en de wettelijke nagroei-indicatoren Koloniegetal bij 22°C (KG22) en *Aeromonas*. Er is geen enkelvoudige relatie gevonden die het verband tussen de ruwwaterkwaliteit, de ontwerp- of bedrijfsvoeringsgerelateerde eigenschappen en nagroei-indicatoren volledig en nauwkeurig bepaalt. Wel zijn verschillende enkelvoudige en meervoudige relaties gevonden die afzonderlijk of gezamenlijk verklarend kunnen zijn voor een deel van de waargenomen variantie van de nagroei-indicatoren. De gevonden verbanden kunnen in gericht vervolgonderzoek toegepast worden om het zuiveringsproces te optimaliseren voor de verwijdering van de klassieke grondwaterparameters en de beperking van de nagroei.



Aeromonas en Koloniegetal bij 22°C (KG22) van het gedistribueerde water bij verschillende beluchtingstypen

Belang: verbetering biologische stabiliteit door extra inzicht in effect van de zuivering

Circa 90% van Nederlandse productielocaties gebruikt grondwater als bron. Bij sommige locaties is groei van KG22 en *Aeromonas* relatief hoog en kan de biologische stabiliteit worden verbeterd. Meer inzicht in het effect dat ontwerp- en bedrijfsvoeringsgerelateerde eigenschappen van de zuivering hebben op de biologische stabiliteit en de chemische waterkwaliteit kan helpen bij die verbetering.

Aanpak: systematische data-analyse op database met gegevens van 43 productielocaties

Het onderzoek is gestart met de selectie van geschikte Nederlandse productielocaties die water bereiden uit anoxisch grondwater. De betrokken waterbedrijven hebben een gestandaardiseerd format ingevuld voor 43 productielocaties, dat is aangevuld met Rewab waterkwaliteitsgegevens en afgeleide eigenschappen. De database bevat gemiddeld 400 eigenschappen per zuivering. De kwaliteit van de database is gecontroleerd en de database is geïmporteerd in de analysesoftware Python. Vanwege de eigenschappen van de database zijn vooral enkelvoudige relaties bestudeerd. In beperkte mate zijn ook hypothese-gedreven meervoudige relaties onderzocht.

Resultaten: omvangrijke, bruikbare database en diverse relaties gevonden

De omvangrijke database is bruikbaar bij dit en ander onderzoek naar de biologische stabiliteit en de verbetering van grondwaterzuivering. Er is geen enkelvoudige relatie gevonden die volledig en nauwkeurig de variantie bepaalt tussen de ruwwaterkwaliteit en ontwerp- of bedrijfsvoeringsgerelateerde eigenschappen enerzijds en nagroei-indicatoren anderzijds. Wel zijn verschillende enkelvoudige en meervoudige relaties gevonden die afzonderlijk of gezamenlijk (deels) verklarend kunnen zijn voor de waargenomen variantie van de nagroei-indicatoren. De belangrijkste verbanden zijn:

- Een positieve relatie tussen het organische stofgehalte in het gedistribueerde water en het methaangehalte in ruwwater en de nagroei-indicatoren.

- Sproeibeluchting lijkt minder gunstig voor de productie van biologisch stabiel water. Torenbeluchting (BOT) lijkt gunstiger dan plaatbeluchting voor de productie van biologisch stabiel water.
- Een constante bovenwaterstand lijkt gunstiger dan een oplopende bovenwaterstand voor de productie van biologisch stabiel water.
- Drukfiltratie lijkt gunstiger dan gravitatiefiltratie voor de productie van biologisch stabiel water.
- Er is een positief verband gevonden tussen de lucht-waterverhouding en de nagroei-indicatoren in het gedistribueerde water.
- Er is een positief verband gevonden tussen de bovenwater contacttijd van de filters en KG22 in het gedistribueerde water.
- Er is een positief verband gevonden tussen het spoelwaterpercentage en KG22 van het gedistribueerde water.
- Er zijn aanwijzingen gevonden voor een mogelijk negatief effect van biologische ontijzering op de biologische stabiliteit van drinkwater. Voor de locaties waarbij de ontijzering mogelijk biologisch van aard is, geldt dat de locaties met een lage lucht-waterverhouding beter presteren.
- Filterschoonmaak, eerste filtraat afvoer en de afwezigheid van spoelwaterhergebruik lijken gunstig voor KG22 in het gedistribueerde water.

Implementatie: nieuwe leads voor relaties tussen eigenschappen zuivering en biologische stabiliteit

Er zijn verschillende enkelvoudige en meervoudige relaties gevonden die afzonderlijk of gezamenlijk (deels) verklarend kunnen zijn voor de waargenomen variantie van de nagroei-indicatoren. De gevonden verbanden kunnen in gericht vervolgonderzoek gebruikt worden om het zuiveringsproces te optimaliseren in relatie tot de verwijdering van de klassieke grondwaterparameters en de beperking van de nagroei.

Rapport

Dit onderzoek is beschreven in rapport *Relatie tussen ontwerp en bedrijfsvoering van grondwaterzuivering en biologische activiteit* (BTO 2016.061).

Samenvatting

Achtergrond en projectdoelstelling

Circa negentig procent van de Nederlandse drinkwaterproductielocaties gebruikt grondwater als bron. Ongeveer twee derde van het Nederlandse drinkwater wordt uit grondwater bereid. In het algemeen wordt grondwater gezuiverd middels één of meerdere beluchting en snelfiltratie stappen. Het doel van de waterbehandeling is voornamelijk het verwijderen van methaan, ammonium, ijzer en mangaan. In de zuiveringsprocessen worden deeltjes en biologisch afbreekbaar organisch koolstof gevormd en verwijderd, tevens wordt een deel van het totaal organische koolstof gehalte verwijderd. De zuiveringsprocessen hebben verschillende invloeden op de biologische activiteit in het drinkwater. Deze invloeden zijn afhankelijk van de ruwwaterkwaliteit, de procesconfiguratie, het procesontwerp en de bedrijfsvoering van de zuivering. In het distributienet kan de biologische activiteit in het water beïnvloed worden door het leidingmateriaal, leidingnetontwerp en de verblijftijd.

Stoffen die aanwezig zijn in ruwwater worden volgens verschillende principes verwijderd (fysisch, chemisch of biologisch). Wanneer de omzetting biologisch van aard is, ontstaat microbiële biomassa in de zuivering. Onduidelijk is of in de processen van de klassieke grondwaterzuivering vorming van makkelijk afbreekbaar koolstof (AOC) plaatsvindt. Als de zuivering niet optimaal functioneert, kan het reinwater lage concentraties van in het ruwwater aanwezige stoffen en/of biomassa bevatten. De doorslag van deze opgeloste en/of deeltjesgebonden stoffen naar het reinwater kan resulteren in biologisch minder stabiel drinkwater en verhoogde groei in het distributiesysteem. In eerder onderzoek is aangetoond dat de aanwezigheid van bepaalde opgeloste en deeltjesgebonden stoffen een negatieve invloed kunnen hebben op de biologische stabiliteit van reinwater en bijdragen aan groei in het leidingnet. Deze stoffen kunnen dienen als voedingsstoffen voor micro-organismen en de deeltjes bieden oppervlak voor de groei van micro-organismen. De ruw- en reinwaterkwaliteit kunnen de waarden van de wettelijke groei-indicatoren (koloniegetal bij 22°C en koloniegetal *Aeromonas*) deels verklaren. In dit onderzoek is het effect van het ontwerp en de bedrijfsvoering van de beluchting en snelfiltratie in grondwaterzuivering betrokken door de eigenschappen van een groot aantal zuiveringen systematisch te analyseren.

In de voorliggende studie is onderzocht of verbanden aanwezig zijn tussen de ruwwaterkwaliteit, het ontwerp en de bedrijfsvoering van de beluchting- en filtratieprocessen van grondwaterproductielocaties die anoxisch grondwater behandelen en de aanwezigheid of verwijdering van typische grondwater parameters, zoals methaan, ammonium, organisch koolstof, ijzer, mangaan, en de troebelings, én de groeiparameters Koloniegetal bij 22°C (KG22) en *Aeromonas* van het geproduceerde reinwater en het gedistribueerde drinkwater. Het onderzoek richt zich op productielocaties die anoxisch grondwater behandelen, omdat het drinkwater bereid uit oxisch grondwater meestal een zeer hoge biologische stabiliteit heeft en weinig groei laat zien. De mogelijke verbanden die uit deze inventarisatie volgen, kunnen in vervolgonderzoek gebruikt worden om het optimale zuiveringsproces voor de behandeling van anoxisch grondwater in relatie tot de verwijdering van klassieke grondwaterparameters vast te stellen. Daarmee kan de biologische stabiliteit van het drinkwater verder worden verbeterd, wat zal leiden tot een vermindering van groei in het distributiesysteem.

Aanpak database

Op basis van een selectie van geschikte productielocaties die anoxisch grondwater gebruiken voor de productie van drinkwater zijn voor 43 productielocaties ontwerp-, bedrijfsvoering en waterkwaliteit gerelateerde gegevens door de drinkwaterbedrijven of anderszins aangeleverd. Het daartoe opgestelde format bevat meer dan duizend ontwerp-, bedrijfsvoering- en waterkwaliteit gerelateerde inputmogelijkheden. Voor elke productielocatie zijn gemiddeld ca. 400 eigenschappen (attributen) beschikbaar, met een minimum van 225 en een maximum van 680, waarvan ca. 70 – 85 % ontwerp- of bedrijfsvoeringsgerelateerde eigenschappen zijn en het resterend deel zijn waterkwaliteit parameters.

Voorafgaand aan de data-analyse is de kwaliteit van de database verbeterd door de aangeleverde informatie in het gestandaardiseerd format te controleren. Er is een grote inspanning geleverd om de database voor te bereiden, vanwege (i) de noodzakelijke kwaliteitscontrole en consistentiechecks, (ii) het combineren van de informatie afkomstig van verschillende bronnen, (iii) het reorganiseren van alle afzonderlijke formats tot één overall database op zodanige wijze dat deze geïmporteerd kan worden in de analyse software (Python). Overige belangrijke bewerkingsstappen van de database zijn (iv) het toevoegen van afgeleide eigenschappen (berekend op basis van beschikbare input), en (v) het aggregeren van informatie op beluchtingsstap- of filterniveau tot locatieniveau, wat noodzakelijk is om productielocaties met verschillende procesconfiguraties (verschillende typen en aantallen procesunits, die parallel, in serie, samenkomend en/of splitsend geschakeld zijn) onderling vergelijkbaar te maken.

Aanpak en conclusie data-analyse

De dataset laat zich als volgt kenmerken: (i) er is heterogene variëteit in de attributen (verschillende procesconfiguraties), (ii) er zijn hiërarchische eigenschappen (eigenschappen kunnen slechts een waarde hebben als een voorgaande eigenschap een waarde heeft), (iii) de matrix is vanwege verschillende redenen 'dun' gevuld (de matrix is *sparse*), (iv) er zijn relatief veel bekende eigenschappen per zuivering in relatie tot het aantal zuiveringen. Door de grote variëteit aan informatie en een relatief klein aantal onafhankelijke observaties wordt een geautomatiseerde data-analyse in Python bemoeilijkt. Gezien deze kenmerken is het trainen van complexere, *machine-learning* regressiemodellen niet de aangewezen analyseroute en is een alternatieve aanpak gevolgd:

- De Spearman correlatiematrix is voor circa 23.000 enkelvoudige relaties berekend, en na reductie voor circa 11.000 gevisualiseerd in een *correlation heatmap*. De Spearman correlatiecoëfficiënt geeft met een getal tussen de -1 en +1 weer hoe sterk de variabelen met elkaar correleren. Visualisatie in een scatterplot is uitgevoerd voor een beperkt aantal relaties met een sterkere correlatiecoëfficiënt en praktische relevantie.
- Een beperkt aantal hypothese gedreven meervoudige analyses zijn uitgevoerd waarbij mogelijke verbanden met behulp van een 3D of 4D visualisatie onderzocht zijn.
- Een automatische clusteringsanalyse is uitgevoerd om datapunten met soortgelijke eigenschappen te groeperen.

Hiertoe is een *machine-learning* toolbox verder uitgewerkt in de programmeertaal Python. De toolbox heeft de volgende functionaliteit: importeren data uit bronbestanden, afgeleide eigenschappen berekenen en toevoegen, en voorbereiding van de dataset (o.a. verwijdering van eigenschappen met incomplete informatie), regressie-analyse en data-visualisatie. Geconcludeerd wordt dat de database in de huidige vorm een bruikbare bron van informatie is die ingezet kan worden bij onderzoek naar de biologische stabiliteit van water, als naslagwerk, om de prestatie en eigenschappen van overige zuiveringen te vergelijken, en

naar verwachting ook bij onderzoek, ontwerp en optimalisatie van grondwaterzuiveringen in bredere zin.

Conclusie overall projectdoelstelling

Er is geen enkelvoudige relatie gevonden die de variantie volledig en nauwkeurig bepaalt tussen de ruwwaterkwaliteit, de ontwerp- of bedrijfsvoeringsgerelateerde eigenschappen enerzijds en biologische groei indicatoren anderzijds. Wel zijn verschillende enkelvoudige relaties gevonden die afzonderlijk of gezamenlijk verklarend kunnen zijn voor een deel van de waargenomen variantie van de groei indicatoren. Daarnaast is een beperkt aantal meervoudige analyses (met drie of vier parameters) gericht op "handmatige" gekozen hypothesen op kwalitatieve wijze onderzocht. Hieruit volgen relaties die afzonderlijk of gezamenlijk (deels) verklarend kunnen zijn voor de waarde van de groei indicatoren. Deze methode is tijdrovend, maar lijkt geschikt om op hoofdlijnen inzicht te verkrijgen in de samenhang tussen drie of vier parameters.

In dit project is niet onderzocht of in de betreffende gevallen sprake is van een causaal verband. Sommige gevonden relaties voldoen aan het verwachtingspatroon, andere relaties geven een onverwacht beeld. Voor de gevonden verbanden zijn mogelijk achterliggende oorzaken geformuleerd, deze kunnen in gerichte vervolgstudies onderzocht worden.

Relaties groei indicatoren

Hierna volgt een overzicht van de verbanden die in dit onderzoek zijn gevonden tussen de waarden van groei indicatoren en de ontwerp- en bedrijfsvoeringsgerelateerde eigenschappen van de zuivering en overige waterkwaliteit parameters van grondwater productielocaties. De relaties tussen een groot aantal eigenschappen is onderzocht. Voor een aantal parameters is een relatie met de groei-indicatoren KG22 en *Aeromonas* gevonden, wat een eerste aanwijzing vormt dat deze parameters de mate van groei mogelijk beïnvloeden. De gevonden verbanden verklaren niet 100% van de variantie, echter zij geven mogelijke richtingen aan voor de optimalisatie van productielocaties en gericht vervolg onderzoek. De aanwijzingen zijn gebruikt om hypothesen op te stellen waarmee de invloed van deze parameters op de groei-indicatoren gericht kan worden onderzocht.

Waterkwaliteit

Het TOC gehalte in het gedistribueerde water en het methaan gehalte in ruwwater tonen een positieve relatie met de groei indicatoren. De verbanden tussen het TOC gehalte van het reinwater en het ammonium gehalte van het ruwwater enerzijds en de groei indicatoren anderzijds zijn in deze studie minder significant dan in eerder BTO onderzoek (van der Wielen, 2011).

Hypothese ontwerp zuivering

- Type beluchting. Sproeibeluchting lijkt minder gunstig voor de productie van biologisch stabiel water. BOT beluchting lijkt gunstiger dan plaatbeluchting voor de productie van biologisch stabiel water.
- Bovenwaterstand. Een constante bovenwaterstand lijkt gunstiger dan een oplopende bovenwaterstand voor de productie van biologisch stabiel water.
- Type filtratie. Drukfiltratie lijkt gunstiger dan gravitatiefiltratie voor de productie van biologisch stabiel water.

Relaties bedrijfsvoering zuivering

- Lucht-water verhouding. Er is een positief verband gevonden tussen de lucht-water verhouding en de waarde van groei indicatoren in het gedistribueerde water.

- Bovenwater contacttijd. Er is een positief verband gevonden tussen de bovenwater contacttijd van de filters en de waarde van het KG22 in het gedistribueerde water.
- Filterspoeling. Er is een positief verband gevonden tussen het spoelwaterpercentage en het genormaliseerde KG22 van het gedistribueerde water. Een vergelijkbaar verband is gevonden tussen het genormaliseerd spoelluchtgetal en het genormaliseerd samengesteld nagroei getal.
- Variabiliteit. Er is een positief verband gevonden tussen het aantal parallelle filterunits (een maat voor het uitdempen van variaties in de zuivering) en de nagroei indicatoren. Er is geen relatie gevonden tussen de waarden van nagroei indicatoren en de variabiliteit van de bedrijfsvoering gebaseerd op de wijze van inbedrijfname van een filter (direct of geleidelijk), de schakelfrequentie van filters, de schakelfrequentie van pompputten en de mate van variatie van productiedebieten.
- Biologische ontijzering. Er zijn aanwijzingen gevonden voor mogelijk negatief effect van biologische ontijzering op de biologische stabiliteit van drinkwater. Voor de locaties waarbij de ontijzering mogelijk biologisch van aard is geldt dat de locaties met een lage lucht-water verhouding beter presteren.
- Optimale filtratiesnelheid. Er is geen verband gevonden tussen de filtratiesnelheid en nagroei-indicatoren, wat duidt op het afwezig zijn van een optimale filtratiesnelheid.
- Bedrijfsvoeringsparameter met binaire waarde. Filter schoonmaak, eerste filtraat afvoer, en de afwezigheid van spoelwaterhergebruik lijken gunstig voor het KG22 in het gedistribueerde water.

Relaties chemische en fysische parameters

De verlaging van het samengesteld chemisch getal (conversie ijzer, mangaan, ammonium) heeft een positieve relatie met de filterbed contacttijd en het spoelwaterpercentage, en een negatieve relatie met de filtratiesnelheid. Vergelijkbare verbanden zijn ook gevonden voor de conversie van de afzonderlijke componenten (ijzer, mangaan, ammonium) en TOC. De hogere ijzergehaltes in reinwater (tot 0,05 mg/L) komen voor bij de productielocaties die een laag ijzergehalte in het ruwwater hebben.

Overige verbanden

Met de onderzochte database zijn geen verbanden gevonden tussen de nagroei indicatoren en de overige ontwerpgerelateerde en bedrijfsvoeringsgerelateerde eigenschappen. Een volledig overzicht van de ontwerpgerelateerde eigenschappen is in Bijlage I weergegeven, de bijhorende Spearman correlatie coëfficiënten zijn beschikbaar maar in dit rapport niet weergegeven. De afwezigheid van (één op één) relaties hoeft niet te betekenen dat parameters definitief geen invloed hebben op nagroei indicatoren. De relatie kan door verstoringen door andere parameter(s) zijn weggevallen of is in combinatie met andere parameters (meervoudige regressie) wel van belang.

Aanbevelingen

De database is in de huidige vorm een bruikbare bron van informatie, in te zetten bij onderzoek naar de biologische stabiliteit van water, als naslagwerk, en bij onderzoek, ontwerp en optimalisatie van grondwaterzuiveringen in bredere zin.¹ De kwaliteit van de database kan op verschillende manieren verder verbeterd worden:

- De kwaliteit van de informatie in de database kan verder verbeterd worden met behulp van afstemming met de drinkwaterbedrijven.

¹ In een BTO onderzoeksvoorstel over ontijzering van de Themagroep 'Drinkwatertechnologie van de Toekomst' is voorgesteld om de database te gebruiken. Het voorstel is door het CO goedgekeurd en zal naar verwachting in 2016 starten.

- De database kan verder gecompleteerd worden door ontbrekende relevante eigenschappen te achterhalen in overleg met de drinkwaterbedrijven.
- De database kan uitgebreid worden met nieuwe productielocaties.
- De database kan uitgebreid worden met nieuwe (afgeleide) parameters, zoals de filterbelasting (vracht per m² of m³ filterbed), het fosfaatgehalte of het AOC gehalte.

Uitbreiding van de database, door toevoeging van productielocaties dan wel door toevoeging van ontbrekende gegevens, is aanbevelingswaardig (i) op basis van de bevinding dat de significantie van gevonden correlaties in voorliggend en eerder BTO onderzoek ondanks een redelijke groepsgrootte (respectievelijk 43 en 38) afhangt van de set van productielocaties die in de studie betrokken zijn of de periode waaruit de data afkomstig zijn, en (ii) omdat de mogelijkheden voor het uitvoeren van meervoudige analyses daarmee toenemen.

In het kader van de verbetering van de biologische stabiliteit kan vervolg onderzoek zich richten op de volgende onderdelen:

- Gericht onderzoek op laboratorium-, proefinstallatie- en fullscale schaal naar de vastgestelde hypothesen die zijn opgesteld aan de hand van de vastgestelde relaties tussen ruwwaterkwaliteit, ontwerp zuivering, bedrijfsvoering zuivering en de nagroei-indicatoren.
- De eigenschappen van productielocaties die uit verschillende zuiveringsunits (in serie en/of parallel) zijn geaggregeerd tot locatieniveau. Om het effect van de onzekerheid die met deze aggregatie gepaard gaat te elimineren kan de correlatie analyses uitgevoerd worden voor een selectie van de productielocaties, namelijk diegene die uit één straat of één unit bestaan.
- Bij verschillende productielocaties is sprake van een verstoring effect door de aanwezigheid van zuiveringsstappen anders dan beluchting en filtratie of door bijmenging van water van een andere locatie. Om het effect van de onzekerheid die met deze verstoring gepaard gaat te elimineren kan de correlatie analyses uitgevoerd worden voor een selectie van de productielocaties, namelijk diegene waarbij geen sprake is van een verstoring.
- In voorliggend onderzoek is een beperkt aantal hypothesen op “handmatige” wijze vastgesteld dan wel arbitrair geselecteerd, en deze zijn middels een meervoudige analyse (met drie of vier parameters) op kwalitatieve manier getoetst. Deze methode is tijdrovend en derhalve is het aantal geformuleerde en geteste hypothesen in voorliggend onderzoek niet uitputtend. Een statistisch onderbouwde meervoudige regressieanalyse kan handmatig worden uitgevoerd met SPSS, maar een dergelijke studie zal relatief veel tijd kosten. Softwarematige automatisering van een uitputtende meervoudige analyse, door het testen van alle mogelijke of gewenste combinaties van drie of vier parameters, is mogelijk met behulp van logistische regressie en parameter optimalisatie routine. Vanwege de eigenschappen van de database (*sparse* matrix en grote verhouding tussen eigenschappen en aantal locaties) is een belangrijke randvoorwaarde bij die analyse dat het aantal parameters in de regressie gemaximeerd wordt tot drie of vier. De betrouwbaarheid van de regressie zou getoetst kunnen worden door validatie van het model na toevoeging van een nieuwe locatie. *Principal component analysis* met een beperking van aantal attributen is een tweede mogelijkheid om meervoudige regressie uit te voeren.
- Toepassing van de database door (nieuwe) grondwater productielocaties met hogere waarden van de nagroei indicatoren “handmatig” te vergelijken met de prestaties, het ontwerp en de bedrijfsvoering van vergelijkbare zuiveringen.

Inhoud

Samenvatting	1
Achtergrond en projectdoelstelling	1
Aanpak database	2
Aanpak en conclusie data-analyse	2
Conclusie overall projectdoelstelling	3
Relaties nagroei indicatoren	3
Aanbevelingen	4
Inhoud	7
1 Inleiding	9
1.1 Achtergrond	9
1.2 Doelstelling	10
1.3 Leeswijzer	10
2 Projectaanpak	11
2.1 Inleiding	11
2.2 Selectie productielocaties	11
2.3 Verkrijgen data	16
2.4 Type data en eigenschappen productielocaties	16
2.5 Data voorbereiding	21
2.6 Data-analyse	23
3 Resultaten	27
3.1 Inleiding	27
3.2 Relaties met chemische waterkwaliteit - vergelijk BTO 2011 onderzoek	27
3.3 Relaties nagroei indicatoren en chemische waterkwaliteit	29
3.4 Relaties tussen nagroei indicatoren en ontwerp zuivering	32
3.5 Relaties tussen nagroei indicatoren en bedrijfsvoering zuivering	36
3.6 Hypothese-gedreven relaties nagroei indicatoren	42
3.7 Test invloed binaire variabelen op waterkwaliteit	54
3.8 Relaties chemische waterkwaliteit en ontwerp of bedrijfsvoering	57
3.9 Resultaten clustering	60
4 Gevonden relaties en mogelijke hypothesen	61
4.1 Relaties met nagroei-indicatoren	61
4.2 Relaties met chemische parameters	64
5 Conclusies en aanbevelingen	67
5.1 Conclusies	67
5.2 Aanbevelingen	69

6	Referenties	71
6.1	Referenties	71
Bijlage I Lijst ontwerp en bedrijfsvoeringsgerelateerde eigenschappen		73
Bijlage II Additionele relaties		83
	Relaties ruwwater en bedrijfsvoering	83
	Relaties chemisch-fysische parameters ruwwater en conversie	93
Bijlage III Uitgangspunten data bewerking		95
	In deze bijlage staan verschillende uitgangspunten opgesomd die gehanteerd zijn bij de kwaliteitscontrole en de verwerking van de aangeleverde formats.	95
	Waterkwaliteit	95
	Bedrijfsvoering	95
	Ontwerp	96

1 Inleiding

1.1 Achtergrond

Circa negentig procent van de Nederlandse drinkwaterproductielocaties gebruikt grondwater als bron. Ongeveer twee derde van het Nederlandse drinkwater wordt uit grondwater bereid. In het algemeen wordt grondwater gezuiverd middels één of meerdere beluchting en snelfiltratie stappen. Het doel van de waterbehandeling is voornamelijk het verwijderen van methaan, ammonium, ijzer en mangaan. In de zuiveringsprocessen worden deeltjes en biologisch afbreekbaar organisch koolstof gevormd en verwijderd, tevens wordt het totaal organische koolstof gehalte verwijderd. De zuiveringsprocessen hebben verschillende invloeden op de biologische activiteit in het drinkwater. Deze invloeden zijn afhankelijk van de ruwwaterkwaliteit, de procesconfiguratie, het procesontwerp en de bedrijfsvoering van de zuivering. In het distributienet kan de biologische activiteit in het water beïnvloed worden door het leidingmateriaal, leidingnetontwerp en de verblijftijd.

Stoffen die aanwezig zijn in ruwwater worden volgens verschillende principes verwijderd. Koolstofdioxide en methaan worden fysisch verwijderd. Biologisch afbreekbaar organisch koolstof, niet fysisch verwijderd methaan, ammonium, ijzer en mangaan worden chemisch of biologisch omgezet. Wanneer de omzetting biologisch van aard is, ontstaat microbiële biomassa in de zuivering. Van bepaalde oxidatieprocessen zoals ozon en UV/H₂O₂ is bekend dat assimileerbaar organisch koolstof (AOC) gevormd wordt uit organische stof (NOM). Onduidelijk is of in de processen van de klassieke grondwaterzuivering ook AOC-vorming plaatsvindt. Als de zuivering niet optimaal functioneert, kan het reinwater lage concentraties van in het ruwwater aanwezige stoffen en/of biomassa bevatten. De doorslag van deze opgeloste en/of deeltjesgebonden stoffen naar het reinwater kan resulteren in biologisch minder stabiel drinkwater en verhoogde nagroei in het distributiesysteem.

In de afgelopen jaren zijn onderzoeken uitgevoerd naar de invloed van opgeloste en deeltjesgebonden stoffen op de biologische stabiliteit en nagroei. Aangevoerd is dat de aanwezigheid van de betreffende opgeloste en deeltjesgebonden stoffen een negatieve invloed kunnen hebben op de biologische stabiliteit van reinwater en bijdragen aan nagroei in het leidingnet. Deze stoffen kunnen dienen als voedingsstoffen voor micro-organismen en de deeltjes bieden oppervlak voor de groei van micro-organismen (Van der Kooij et al, 2012; Van der Kooij et al, 2013; Van der Kooij et al, 2014, Van der Kooij & Bereschenko, 2015).

In eerder BTO-onderzoek is een selectie van grondwater productielocaties op basis van hun ruwwatersamenstelling ingedeeld, is de mate van nagroei, gebaseerd op het koloniegetal bij 22°C (KG22) en *Aeromonas*, in het distributiesysteem gerangschikt en zijn mogelijke verschillen getracht te verklaren aan de hand van procesontwerp en bedrijfsvoering (van der Wielen, P. et al, 2011, Van Leerdam, R., 2012). Daarnaast heeft RoyalHaskoningDHV onderzocht welke relaties bestaan tussen de ruwwaterkwaliteit en het procesontwerp en de bedrijfsvoering van de beluchting en filtratie van een selectie van grondwater productielocaties. In de betreffende filterdatabase zijn de productielocaties van Vitens, WML, Evides en Oasen betrokken. De resultaten van die studie lieten onder meer zien welke verschillen bestaan tussen en binnen bedrijven in gebruik van het type beluchting (bijvoorbeeld cascade-, toren-, sproei- en plaatbeluchting), filterontwerp (bijvoorbeeld bedhoogte), filtermateriaal (bijvoorbeeld type en grootte) en de bedrijfsvoering (bijvoorbeeld speelregime filters) en de relatie tussen de ruwwaterkwaliteit en deze eigenschappen.

Waterbedrijven voeren geregeld onderzoek uit naar de relatie tussen ontwerp en bedrijfsvoering van een installatie en de reinwaterkwaliteit. Vaak zijn die studies gericht op één of enkele productielocaties. Een systematische studie naar het effect van de ontwerp- en bedrijfsvoeringsverschillen van beluchting en filtratie op de verwijdering, omzetting en vorming van stoffen die karakteristiek zijn voor grondwater en de biologische stabiliteit of nagroei is niet eerder uitgevoerd voor een brede selectie van Nederlandse productielocaties die anoxisch grondwater behandelen met (tenminste) beluchting en snelfiltratie.

De ruw- en reinwaterkwaliteit kunnen de waarden van de wettelijke nagroei-indicatoren (koloniegetal bij 22°C en koloniegetal *Aeromonas*) deels verklaren (van der Wielen, 2011). In dit onderzoek is bekeken in hoeverre de waarden van de nagroei indicatoren daarnaast op basis van de zuiveringsgerelateerde eigenschappen verklaard kunnen worden.

1.2 Doelstelling

In dit project is gezocht naar verbanden tussen de ruwwaterkwaliteit, het ontwerp en de bedrijfsvoering van de beluchting- en filtratieprocessen van grondwaterproductielocaties die anoxisch grondwater behandelen en de aanwezigheid of verwijdering van typische grondwater parameters, zoals methaan, ammonium, organisch koolstof, ijzer, mangaan, en de troebeling, én de nagroeiparameters Koloniegetal bij 22°C (KG22) en koloniegetal *Aeromonas* (*Aeromonas*) van het geproduceerde reinwater en het gedistribueerde drinkwater. Het onderzoek richt zich op productielocaties die anoxisch grondwater behandelen, omdat het drinkwater bereid uit oxisch grondwater meestal een zeer hoge biologische stabiliteit heeft en weinig nagroei laat zien.

De mogelijke verbanden die uit deze inventarisatie volgen kunnen in vervolgonderzoek gebruikt worden om het optimale zuiveringsproces voor de behandeling van anoxisch grondwater in relatie tot de verwijdering van klassieke grondwaterparameters vast te stellen. Daarmee kan de biologische stabiliteit van het drinkwater verder worden verbeterd, wat zal leiden tot een vermindering van nagroei in het distributiesysteem.

Het bestaan van een verband tussen ruwwaterkwaliteit, procesconfiguratie, ontwerp en bedrijfsvoering enerzijds, en de reinwaterkwaliteit en het gedistribueerde water anderzijds is een gegeven. De eigenschappen van de zuivering zijn van invloed op de verandering van waterkwaliteit tijdens de zuivering. Onderzocht is of eventuele verschillende prestaties van productielocaties verklaard kunnen worden aan de hand van (combinaties van) de procesconfiguratie, het procesontwerp en de bedrijfsvoering. De 'prestatie' heeft in deze betrekking op een grotere verwijdering en/of een lager absoluut gehalte van ijzer, mangaan, ammonium, methaan dan wel lagere aantallen van KG22 en *Aeromonas*. Deze prestaties zijn vergeleken voor productielocaties met een vergelijkbare ruwwaterkwaliteit en voor productielocaties met vergelijkbare zuiveringseigenschappen.

1.3 Leeswijzer

In hoofdstuk 2 is de projectaanpak beschreven. In hoofdstuk 3 zijn de belangrijkste resultaten gepresenteerd en bediscussieerd. Hoofdstuk 4 presenteert de conclusies en aanbevelingen.

2 Projectaanpak

2.1 Inleiding

In dit hoofdstuk is de projectaanpak omschreven. Achtereenvolgens is omschreven welke Nederlandse productielocaties in de analyse betrokken zijn, op welke wijze de relevante waterkwaliteit-, ontwerp- en bedrijfsvoeringsgerelateerde eigenschappen van die locaties verkregen zijn, en is beschreven op welke wijzen naar verbanden en relaties tussen deze eigenschappen gezocht is.

2.2 Selectie productielocaties

Er zijn meer dan 200 grondwaterproductielocaties in Nederland en circa 180 productielocaties voldoen aan de kwalificatie 'anoxisch'. Middels een selectiefase is het aantal onderzochte productielocaties gereduceerd tot 54. De selectie is aan de hand van de volgende criteria uitgevoerd.

- *ABIKOU indeling.* Er zijn circa 180 productielocaties die volgens de ABIKOU indeling geïnclassificeerd worden met B ((semi)-confined), AB (tussenklasse), I (*artificial recharge and recovery*) en U (oever infiltratie). Naast deze ABIKOU klassen is sporadisch een A (freatisch) winning in de eerste selectie opgenomen. Deze locaties komen in aanmerking voor dit onderzoek. Water met classificatie K (kalksteen) en O (oppervlaktewater) zijn uitgesloten. Aanvullende selectiecriteria zijn noodzakelijk om het aantal locaties te reduceren.
- *Ruwwaterclassificatie.* De productielocaties zijn onder meer geselecteerd op basis van verschillende combinaties van concentraties ammonium en organisch koolstof in ruwwater, zodanig dat het gehele spectrum dat gevormd wordt door combinaties van ammonium en organisch stof gehalte in het ruwwater in de database aanwezig is. Zo zijn ook enkele locaties met een laag ammonium en organisch koolstof gehalte geselecteerd, ondanks de verwachting dat deze locaties drinkwater met een hoge biologische stabiliteit produceren en distribueren ongeacht het ontwerp of de bedrijfsvoering van de zuivering.² Op basis van een beschikbaar overzicht uit eerder BTO onderzoek, met een beperkt aantal ruwwater parameters voor een groot aantal productielocaties, is het aantal productielocaties ingedeeld. Op basis van combinaties van het ammoniumgehalte en het gehalte organisch stof zijn negen³ groepen gedefinieerd en is elke productielocatie in één van die groepen ingedeeld.⁴ De classificatie is gemaakt op basis van een pragmatische keuze voor de waarden van ammonium en organisch koolstof, zoals weergegeven in Tabel 1.

² Gebaseerd op het uitgangspunt dat deze productielocaties volgens dusdanige ontwerpvoorschriften zijn gebouwd en volgens dusdanige bedrijfsvoeringseisen worden bedreven dat de productie van biologisch stabiel water waarschijnlijk is.

³ In beginsel werd nog een extra kwalificatie toegevoegd (een Tussengroep tussen M en L), in de laatste selectiestap zijn deze twee groepen samengevoegd en als één groep 'laag' beschouwd.

⁴ Voor Brabant Water waren in deze fase geen TOC of DOC ruwwater data beschikbaar en is voor het gemak aangenomen dat het organisch stof gehalte in het ruwwater gelijk is aan het organisch stof gehalte in het reinwater

TABEL 1. CLASSIFICATIE RUWWATER OP BASIS VAN AMMONIUM EN ORGANISCH KOOLSTOF GEHALTE.

Classificatie	Ammonium [mg/L]	Organisch koolstof (TOC of DOC) [mg/L]
L (laag)	< 0,1	< 1,0
M (midden)	0,1 – 0,5	1,0 – 4,0
H (hoog)	> 0,5	> 4,0

- *Overall score nagroei.* Voor een aantal voorzieningsgebieden is in eerder BTO onderzoek een “overall scores voor nagroei” gepresenteerd op basis van *Aeromonas* en KG22, voor drinkwater en gedistribueerd water, uitgedrukt in een getal tussen 1 – 12 (BTO2011.001, BTO2014.206(s)). Deze scores zijn niet voor alle afzonderlijke locaties beschikbaar waardoor de selectie niet alleen op basis van deze score gemaakt kan worden. Bij de keuze tussen gelijkwaardige locaties is deze overall score bij het selecteren betrokken
- *Waterbedrijven.* Bij voorkeur worden productielocaties geselecteerd van alle waterbedrijven die grondwater behandelen.
- *Zuivering.* Bij voorkeur worden zuiveringsinstallaties geselecteerd die enkel bestaan uit (combinaties van) beluchting en snelfiltratie. Ten behoeve van dit selectie criterium is gebruik gemaakt van beschikbare zuiveringsgegevens van Waterbedrijf Groningen, WML, WMD, Oasen en Evides, en zijn de *process flow diagrams* van de locaties van Vitens en Brabant Water opgevraagd omdat de productielocaties van de twee laatst genoemde bedrijven het grootste aandeel vormen. Om het aantal productielocaties in de selectie te reduceren zijn verschillende zuiveringsgerelateerde criteria gehanteerd die in Tabel 2 zijn samengevat.

TABEL 2. ZUIVERINGSGERELATEERDE SELECTIECRITERIA.

criterium	Geschiktheid voor selectie
Zuiveringen die enkel (combinaties van) beluchting en snelfiltratie bevatten.	Voorkeur
Zuiveringen die beluchting en marmerfiltratie bevatten. Uitsluiting zal voor een belangrijk deel al plaatsvinden op basis van de grondstofindeling volgens Tabel 1. Freatisch water wordt vaak met marmer behandeld. Dit type water lijkt t.b.v. deze studie minder relevant maar in sommige gevallen bevat zulk water ook kleine hoeveelheden ammonium, organisch stof, ijzer en mangaan en kunnen locaties met marmer op basis van die eigenschappen toch geselecteerd kunnen zijn.	Bij voorkeur niet selecteren
Complexere zuiveringen. Dit zijn zuivering die naast beluchting en snelfiltratie aanvullende stappen bevatten, zoals ontharding, UV, AKF, nanofiltratie, ionenwisseling, of waarbij meerdere zuiveringsstraten aanwezig zijn, of waarbij menging met andere waterstromen (halfabricaat of drinkwater) plaatsvindt. Zulke additionele stappen (t.o.v. beluchting en snelfiltratie) worden beschouwd als een extra, niet gewenste variabele (<i>disturbance</i>) omdat deze een effect hebben op de waterkwaliteit en dat niet als onderzoeksdoel in voorliggend project wordt onderzocht.	Bij voorkeur niet selecteren
Sluiting van locaties. Locaties die op korte termijn gesloten worden kunnen nuttige ontwerp- en bedrijfsvoeringsgerelateerde input leveren. Zuiveringen die gesloten zijn in de periode tussen het opstellen van de gebruikte database van vorig BTO onderzoek.	Toestaan Uitsluiten
In sommige ruwwatergroepen is het aantal geschikte productielocaties op basis van andere selectiecriteria niet toereikend. In dat geval worden complexere zuiveringen geselecteerd. - Ontharding, ionenwisseling, actief kool, UV en nanofiltratie worden bij voorkeur uitgesloten omdat ze beschouwd kunnen worden als niet-klassieke grondwaterzuiveringstechnieken en een effect hebben op biologische stabiliteit. In een aantal gevallen zijn zuiveringen met ontharding, AKF, UV of nanofiltratie geselecteerd. ⁵ Locaties met ionenwisseling (IEX) zijn niet geselecteerd. - Indien zuiveringen met een extra, niet gewenste variabele (<i>disturbance</i>) geselecteerd moeten worden, dan gaat de voorkeur uit naar een minimaal aantal van dat soort ongewenste variabelen. Bij deelstroom behandeling is het effect van de ongewenste variabele afhankelijk van de ratio van de deelstroom, echter deze was in de selectiefase onbekend.	Concessie
Random keuze Wanneer het aantal geschikte productielocaties groter is dan het benodigd aantal, en aan bovenstaande criteria wordt voldaan, heeft de selectie random plaatsgevonden. Dit is in twee gevallen gebeurd (4 random kiezen uit 6, en elders 6 random kiezen uit 15).	Selecteren

⁵ Als er keuze is tussen twee locaties van één bedrijf (een met en een zonder ontharding) gaat de voorkeur uit naar de locatie zonder ontharding.

Sommige productielocaties hebben naast beluchting en filtratie aanvullende zuiveringsstappen. Deze locaties zijn weliswaar in het onderzoek betrokken, en de aanvullende zuiveringsstappen zijn als proces benoemd in de inventarisatie, echter enkel de effecten van het ontwerp en de bedrijfsvoering van beluchting en snelfiltratie op de waterkwaliteit is bestudeerd.

Op basis van bovengenoemde criteria werden 54 productielocaties geselecteerd. De productielocaties die na de selectiefase in het onderzoek betrokken zijn, wijkt hier op twee punten van af: er zijn 43 productielocaties betrokken in de analyse, en een deel van de geselecteerde productielocaties is vervangen door een alternatieve locatie. De reden voor deze afwijking hangt samen met de beschikbare tijdsbesteding van de waterbedrijven die de gegevens over de productielocaties hebben aangeleverd. In Tabel 3 zijn de betrokken productielocaties en de ruwwater classificatie weergegeven.

TABEL 3. ONDERZOCHE PRODUCTIELOCATIES.

Productielocatie	Ruwwaterclassificatie	Waterbedrijf
Roosendaal	MM	Brabant Water
Budel	HM	Brabant Water
Dorst	HM	Brabant Water
Wouw	HH	Brabant Water
Eindhoven	MM	Brabant Water
Genderen	HM	Brabant Water
Son	MM	Brabant Water
Someren	ML	Brabant Water
Helmond	HM	Brabant Water
Huijbergen	MM	Evides
Ossendrecht	HM	Evides
Halsteren	HH	Evides
Lekkerkerk - Winning Tiendweg	HM	Oasen
Lekkerkerk - Winning Schuwacht	HM	Oasen
Holk	ML	Vitens
Espelo	MH	Vitens
Zutphen	HM	Vitens
Eerbeek	ML	Vitens
Goor	HH	Vitens
Veenendaal	ML	Vitens
Herikerberg	LM	Vitens
Vorden	MM	Vitens
Fikkersdries I	LL	Vitens
Fikkersdries II	LL	Vitens
Weerselo	HM	Vitens
Beerschoten	ML	Vitens
Hasselo	HH	Vitens
Diepenveen	HH	Vitens
Leidsche Rijn	MM	Vitens
Zoelen	MM	Vitens
Zutphenseweg	HM	Vitens
Noardburgum	HH	Vitens
De Meern	MM	Vitens
Sellingen	MH	Waterbedrijf Groningen
Onnen	MH	Waterbedrijf Groningen
De Groeve	MM	Waterbedrijf Groningen
Nietap	MH	Waterbedrijf Groningen
Assen	HM	WMD
PS Hoogeveen	HH	WMD
Annen	LM	WMD
Zuidwolde	HH	WMD
Beegden	HH	WML

2.3 Verkrijgen data

In het onderzoek is gebruik gemaakt van waterkwaliteit parameters, ontwerpgegevens van productielocaties en bedrijfsvoeringstechnische eigenschappen van de productielocaties. Deze data zijn op de volgende wijzen verkregen:

- *Waterkwaliteit.* Ruwwater en reinwater gegevens zijn opgehaald uit de Rewab database. De Rewab indeling is in bepaalde gevallen inadequaat voor de koppeling van waterkwaliteitsgegevens aan productielocaties, vooral voor sommige voorzieningsgebieden is koppeling tussen de Rewab distributie data en bijhorende productielocatie niet goed mogelijk. De relevante waterkwaliteitsgegevens van het gedistribueerde water zijn aangeleverd door de drinkwaterbedrijven. Er is gebruik gemaakt van data in de periode 2011 – 2013.
- *Ontwerp- en bedrijfsvoeringseigenschappen.* De ontwerp- en bedrijfsvoeringsgerelateerde eigenschappen zijn door de waterbedrijven aangeleverd. Daartoe is in Excel een format opgesteld. Dit uitgebreide format omvat een breed scala van eigenschappen. Het format heeft het “Filterpaspoort” als uitgangspunt genomen, vervolgens is daar een groot aantal eigenschappen aan toegevoegd. Het format bevat de eigenschappen (waarden in te vullen door de bedrijven), de eenheden van deze eigenschappen, uitleg over de betreffende parameter, en is dusdanig opgezet dat vast staat of de waarde een tekst string, een continue waarde, een voorgeselecteerde tekst string (keuzemenu) of een TRUE/FALSE statement betreft. De waterbedrijven hebben met een omvangrijke inspanning deze formats ingevuld. Er is gekozen om het format in deze uitgebreide vorm toe te passen i) omdat is verondersteld dat detail informatie benodigd is om verbanden te vinden die nog niet eerder zijn aangetoond, ii) om geen ontwerp- en bedrijfsvoeringsgerelateerde eigenschappen op voorhand uit te sluiten, en iii) omdat het wat betreft tijdsbesteding vaak efficiënter is om gegevens bij de bedrijven één maal te verzamelen in plaats van te moeten terugkeren voor aanvullende gegevens.
- *Afgeleide parameters.* Op basis van de door de waterbedrijven aangereikte informatie zijn verschillende nieuwe, afgeleide parameters bepaald (zogenaamde *lazy attributes*). Als voorbeeld kan gedacht worden aan de filtratiesnelheid die op basis van het debiet en het filteroppervlak is bepaald. Het format is hier in tweede instantie op aangepast.

2.4 Type data en eigenschappen productielocaties

In deze paragraaf wordt beschreven met welke gegevens zijn geïnventariseerd en geanalyseerd om de relatie tussen procesontwerp en de bedrijfsvoering van beluchting en snelfiltratie met de verwijdering van stoffen kenmerkend voor grondwater in relatie tot nagroei te doorgronden.

2.4.1 Waterkwaliteit

De volgende waterkwaliteitsgegevens zijn bij de analyse betrokken: ijzer, mangaan, ammonium, methaan, organisch koolstof (TOC of DOC), zuurgraad, troebelingsgraad, temperatuur, KG22 en *Aeromonas*. De onderzochte watertypen zijn het ruwwater, het reinwater en het gedistribueerde drinkwater. Voor de chemische en fysische parameters zijn de minimum, gemiddelde en maximale waarden beschikbaar. In dit onderzoek is gefocust op relaties met de gemiddelde waarden. Voor de microbiologische parameters zijn het minimum, het gemiddelde, de mediaan, het geometrisch gemiddelde en de maximale waarde beschikbaar voor het gedistribueerde water. In dit onderzoek is gefocust op relaties met de geometrische gemiddelden. Voor het reinwater is gebruik gemaakt van de Rewab database, en zijn de minimum, rekenkundig gemiddelde (*Aeromonas*) en geometrisch gemiddelde (KG22) en maximum waarden beschikbaar. In dit onderzoek is gefocust op relaties met de gemiddelde waarden.

De microbiologische parameters KG22 en *Aeromonas* zijn betrokken vanwege het feit dat deze parameters de wettelijke indicatoren voor nagroei vormen. Daarnaast is in het verleden waargenomen dat AOC correleert met KG22 en BVS met *Aeromonas*. Deze parameters zijn daarmee relevant en voor alle productielocaties beschikbaar. Het gebruik van andere indicatoren voor biologische activiteit, zoals de parameters assimileerbaar koolstof (AOC), adenosinetrifosfaat (ATP) of de biofilmvormingspotentie, hebben ten opzichte van de wettelijke indicatoren absoluut meerwaarde of geven zelfs een vollediger beeld, echter deze parameters zijn voor de meeste locaties niet voor handen.

2.4.1.1 Waterkwaliteit data productielocaties

Het gemiddelde ammonium en mangaan gehalte in het reinwater en het gedistribueerde water is vaak gelijk aan de rapportagegrens. Het methaan gehalte in het reinwater wordt vaak niet gemeten of is gelijk aan de rapportagegrens. Correlaties met de absolute waarden van deze parameters geven weinig inzicht, in deze gevallen kunnen de afgeleide parameters, zoals de conversie, een beter beeld geven. Het KG22 en *Aeromonas* van het reinwater zijn in veel gevallen laag: het KG22 is in negen van de 43 gevallen groter dan 5 kve/mL en *Aeromonas* is in acht gevallen groter dan 5 kve/100 mL. De waarden van deze nagroei indicatoren in het gedistribueerde water zijn hoger, daarom zijn de correlaties met die parameters van groter belang.

2.4.1.2 Correlaties BTO onderzoek 2011

In eerder onderzoek is aangetoond is dat de aanwezigheid van de opgeloste en deeltjesgebonden stoffen zoals ijzer, mangaan, organische koolstof en ammonium, een negatieve invloed kunnen hebben op de biologische stabiliteit van reinwater en kunnen bijdragen aan nagroei in het leidingnet. Deze stoffen kunnen dienen als voedingsstoffen voor micro-organismen en de deeltjes bieden oppervlak voor de groei van micro-organismen of zijn, in geval van ijzer en mangaan, in staat negatief geladen organisch materiaal te binden, waardoor dit materiaal langdurig beschikbaar blijft voor micro-organismen. Vanwege deze redenen en vanwege de beschikbaarheid van data is de relatie met deze typische grondwaterparameters opgenomen in het onderzoek.

In van der Wielen et al (2011) zijn de relaties tussen de chemische en fysische ruw- en reinwaterkwaliteit met de nagroei indicatoren onderzocht. In Tabel 4 en Tabel 5 wordt een overzicht gegeven van de belangrijkste resultaten.

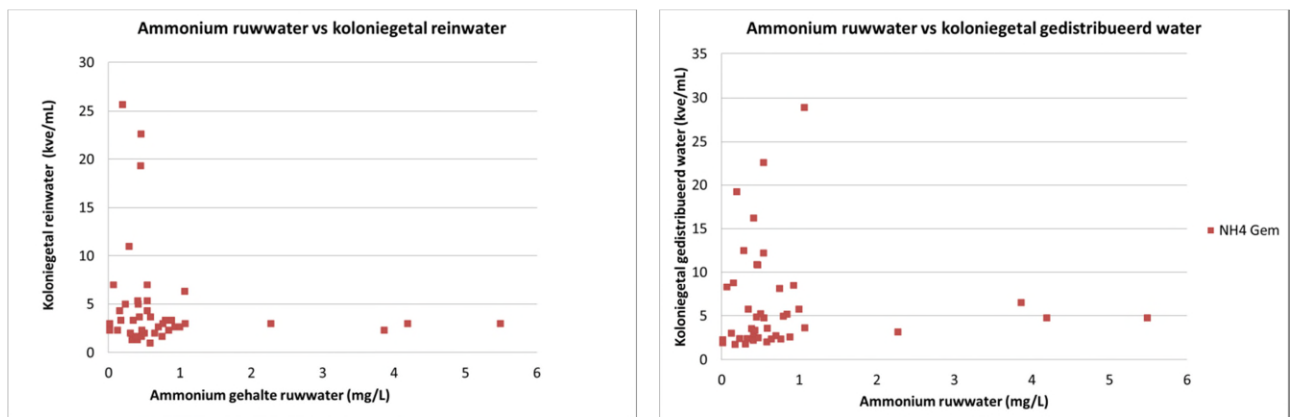
TABEL 4. STATISTISCHE KENMERKEN VAN DE SIGNIFICANTE ENKELVOUDIGE CORRELATIES (PEARSON; $P < 0,01$) TUSSEN HET GEOMETRISCHE GEMIDDELDE VAN KG22 IN HET GEDISTRIBUEERDE DRINKWATER (DISTRIBUTIE) EN FYSISCHE OF CHEMISCHE PARAMETERS IN HET RUWWATER, REINWATER OF GEDISTRIBUEERDE DRINKWATER BIJ 38 GRONDWATERLOCATIES

Parameter	R ²	richting
TOC - ruw	0,59	+
TOC - distributie	0,47	+
TOC - rein	0,47	+
Methaan - ruw	0,41	+
Ijzer - ruw	0,36	+
Ammonium - ruw	0,34	+

TABEL 5. STATISTISCHE KENMERKEN VAN DE SIGNIFICANTE ENKELVOUDIGE CORRELATIES (PEARSON; $P < 0,01$) TUSSEN HET GEOMETRISCHE GEMIDDELDE VAN *AEROMONAS* IN HET GEDISTRIBUEERDE DRINKWATER (DISTRIBUTIE) EN FYSISCHE OF CHEMISCHE PARAMETERS IN HET RUWWATER, REINWATER OF GEDISTRIBUEERDE DRINKWATER BIJ 38 GRONDWATERLOCATIES

Parameter	R^2	richting
TOC - ruw	0,45	+
Ammonium - ruw	0,44	+
TOC - distributie	0,42	+
TOC - rein	0,38	+
Methaan - ruw	0,19	+

Op basis van deze chemische parameters kunnen de waarden van de nagroei indicatoren deels verklaard worden. Anders gezegd: het is niet mogelijk om de waarde van de nagroei indicatoren alleen te verklaren op basis van de ruwwaterkwaliteit. Dit wordt aan de hand van een voorbeeld in Figuur 1 getoond met de data uit voorliggend onderzoek: het blijkt dat sommige productielocaties met een laag ammonium gehalte een hogere waarde van het KG22 hebben, terwijl andere productielocaties met een hoger ammonium gehalte een lage waarde van het KG22 hebben. Een vergelijkbaar beeld geldt voor *Aeromonas* in het gedistribueerde water. Vanwege deze reden is in dit onderzoek gezocht naar aanvullende, zuiveringsgerelateerde verklarende factoren. Tevens is gecontroleerd in hoeverre de hier genoemde correlaties ook in dit onderzoek met een alternatieve set productielocaties gevonden wordt.



FIGUUR 1. AMMONIUM RUWWATER VERSUS KG22 REINWATER (LINKS) EN GEDISTRIBUEERD WATER (RECHTS)

2.4.2 Ontwerp- en bedrijfsvoeringsgerelateerde parameters

Een samenvatting van de ontwerp- en bedrijfsvoeringsgerelateerde eigenschappen die het format bevat is weergegeven in Tabel 6. Een gedetailleerd overzicht van het format is in Bijlage I weergegeven.

TABEL 6. SAMENVATTING FORMAT EIGENSCHAPPEN PRODUCTIELOCATIES.

Categorie	Eigenschap
Algemene informatie productielocatie	Naam waterbedrijf, naam en bouwjaar productielocatie
	Ontwerpcapaciteit, gemiddelde capaciteit. Op jaar-, dag-, uurbasis.
	Procesconfiguratie (<i>process flow diagram</i>)
	Volstroom of deelstroom behandeling, spoelwaterhergebruik, <i>disturbance</i> , ruwwater of tussen kelder.
Beluchtingssectie	Aantal beluchtingsstappen in serie
	Waterdebiet beluchting. Minimaal, gemiddeld, maximaal. Op sectie niveau en beluchter niveau.
	Type beluchting: Cascade, sproeier, beluchting- en ontgassingstoren (BOT), venturi of compressor, plaat, <i>deep well aeration</i> , zuurstofdosing
Cascade	Eigenschappen cascade: Aantal cascades, aantal valstappen, valhoogte, dimensionering bak, type waterverdeling
Sproeier	Eigenschappen sproeibeluchting: Aantal sproeistappen in serie (sproeikamer?), type sproeier (Dresden, Limburg, WMN, Gouds, KSH, sproei-arm, Overig), opwaarts of neerwaarts sproeien, valhoogte water, aantal <i>nozzles</i> .
BOT	Eigenschappen BOT beluchting: Aantal BOTs, type, materiaal en leeftijd pakking, dimensionering, mee/tegenstroom beluchting.
Plaat	Eigenschappen plaatbeluchters: Aantal, dimensionering, dikte waterlaag, aantal en diameter gaten,
Venturi/ compressor	Eigenschappen compressor of venture beluchting: Aantal beluchters, ontluchting
	Zuurstof dosering en <i>deep well aeration</i> : Geen locaties met zuurstofdosing of <i>deep well aeration</i> in de database
	Continuïteit: Beluchting continue in bedrijf, aantal compressoren constant, luchtdebiet regeling
	Overige eigenschappen beluchting: Geforceerd of natuurlijk, luchtdebiet, luchtbehandeling, schoonmaak beluchters, waterkwaliteit effluent beluchting
Filtersectie	Aantal filtratiestappen in serie en parallel. Is het aantal voor- en nafilts gelijk?
	Waterdebieten Minimaal, gemiddeld, maximaal. Op filtersectie niveau en filter niveau. Op uur- en dagbasis.
	Continuïteit: Productiedebiet gedurende een dag (nagenoeg) constant of variabel, filters continu/geschakeld in bedrijf, schakelfrequentie, inbedrijfname geleidelijk of ineens, carousel bedrijf?
	Bovenwaterstand constant of oplopend, (maximale) hoogte bovenwaterstand
Filterbed	Eigenschappen filtermedia en filterbed: Leeftijd, type (grind, marmer, antraciet), deeltjesgrootte, slijtage, aangroei, reiniging. Enkel of dubbellaags, aanwezigheid steunlaag, dimensionering filter (bedhoogte, oppervlak).
	Eigenschappen filters: Gravitatie of druk filtratie, droog of nat filtratie, continue bed beluchting?
Filterspoeling	Eigenschappen spoelwater (ontwerp): Aantal spoelgoten, positie spoelgoten (zijkant/midden en korte/lange kant), afstand bed tot spoelgoot, aantal spoeldoppen/m ² , luchtinbreng, spoelluchtbehandeling.
	Eigenschappen spoelwater (bedrijfsvoering): Spoelwaterprogramma (spoel-% water en lucht), spoelcriterium (tijd, volume, waterkwaliteit, drukval), looptijd, handmatig of automatisch, eerste filtraatvoer, spoelprogramma optimalisatie?

2.4.3 Afgeleide parameters

Verschillende relevante parameters zijn niet direct middels het format aangeleverd maar zijn op basis van de geïnventariseerde data berekend. In tabel is weergegeven om welke parameters het gaat en wat de achterliggende berekening is.

Afgeleide parameter	Berekend op basis van parameters
Filtratiesnelheid	Waterdebiet en filteroppervlak
Maximale spoelsnelheid (water en lucht)	Water- en luchtdebiet tijdens spoeling en filteroppervlak
Spoelwater- en spoellucht volume	Spoelwater- en spoellucht volume op basis van het spoelprogramma (fasen, duur, debieten) genormaliseerd voor het filteroppervlak. Spoelwater- en spoellucht percentage op basis van het spoelprogramma (fasen, duur, debieten), de filterlooptijd, het aantal filters (serie, parallel) en de jaarproductie van de betreffende filtersectie, rekening houdend met deelstroombehandeling.
Benuttingsgraad (of <i>Capacity factor</i>)	Ratio tussen daadwerkelijke productie en ontwerpcapaciteit
Lucht-water verhouding beluchting	Ratio waterdebiet en luchtdebiet in beluchters
Beluchting effectiviteitsgetal	Zie separate toelichting.
Verblijf- en contacttijd	Filters: waterdebiet, filteroppervlak en bedhoogte of bovenwaterstand Beluchting: waterdebiet en dimensionering beluchters. Contacttijd venturi en sproeier: zie separate toelichting
Filterbed contacttijd	Berekend op basis van het debiet, het filterbedvolume (oppervlak x bedhoogte)
Bovenwater contacttijd	Berekend op basis van het debiet en het bovenwatervolume (filteroppervlak x bovenwaterstand). Bij oplopende bovenwaterstand is gerekend met de helft van de maximale bovenwaterstand.
Verwijderingsrendement chemische parameters	Absolute en relatieve conversies van typische grondwater parameters, op basis van de gemiddelde ruw- en reinwater kwaliteit. Per afzonderlijke parameter, en na normalisatie (op basis van de wettelijke normwaarde) samengesteld (som ijzer+mangaan+ammonium/3). Conversies gelijk aan 0% worden verondersteld te worden veroorzaakt door waarden gelijk aan de rapportagegrens, deze zijn buiten de samengestelde parameter gehouden, de noemer is hierop gecorrigeerd.
(Vorming) nagroei indicatoren	Na normalisatie (op basis van de normwaarde voor het KG22 van 100 kve/mL en voor <i>Aeromonas</i> van 1000 kve/100 mL), afzonderlijk getal en samengesteld getal (som genormaliseerde indicatoren/2) voor de vorming van deze microbiologische parameters.
Geaggregeerde parameters	Aggregatie van eigenschappen op unit- of straatniveau tot één kengetal voor de hele locatie. Zie separate toelichting.

2.4.3.1 Beluchting Effectiviteitsgetal

Het effectief beluchtingsgetal is in dit project ontwikkeld om de karakteristiek van een beluchtingsstap te kwantificeren, om verschillende typen beluchtingen met elkaar te kunnen vergelijken, en om het effect van verschillende beluchtingsstappen binnen één locatie te aggregeren tot één beluchtingsgetal voor de gehele locatie. Het effectief beluchtingsgetal van een beluchtingsstap is gebaseerd op de effectiviteit voor de verwijdering van methaan

(Reijnen, 1994), omdat deze component verband houdt met biologische activiteit en omdat de intensiteit van een beluchting bij het ontwerpen van een zuivering vaak wordt afgestemd op het methaan gehalte. Er is gekozen voor een arbitraire schaal van het effectief beluchtingsgetal tussen 1 en 10 waarbij de scores bepaald zijn op basis van Reijnen (1994) en *expert judgement*:

- Plaat- en BOT beluchters zijn effectief voor methaanverwijdering en krijgen score 9.
- Cascade beluchting kan effectief zijn voor methaan verwijdering, de effectiviteit hangt ondermeer samen met het aantal valstappen en de valhoogte per stap. De cascade is gescoord tussen de 4 en de 9, de exacte score is bepaald met een functie afhankelijk van het aantal stappen en de hoogte per stap. De effectiviteit als functie van de hoogte is gebaseerd op van der Helm (1998).
- De effectiviteit van sproeibeluchting voor de verwijdering van methaan hangt ondermeer af van de contacttijd van de druppels en de lucht. De score van de effectiviteit van de sproeier is geschaald tussen 4 – 6. De score is genormaliseerd op basis van de minimale en de maximale valhoogte die in de database beschikbaar is. De productielocatie met de maximale valhoogte (van 4 m) scoort een 6.
- Venturi en compressor beluchting zijn ineffectief voor de verwijdering van methaan en zijn op 1 gescoord.

2.4.3.2 Contacttijd venturi en sproeier

De contacttijd van venturi of compressor beluchting hangt samen met het exacte procesontwerp. Omdat deze gegevens onbekend zijn is uitgegaan van een standaard situatie van een leiding met 0,4 m diameter, een lengte van 10 m, en een debiet van 60 m³/uur. Daarmee komt de contacttijd neer op ca. 1 minuut.

De bepaling van de contacttijd in de sproeier is gebaseerd op de valvergelijking van de afstand als functie van de (val)tijd. Daarbij is aangenomen dat de uittreed snelheid van water 0 m/s is, de luchtweerstand speelt geen rol, en de valhoogte is gelijk is aan de afstand tot aan het bovenwater. De contacttijd (de valtijd) is gelijk aan $\sqrt{(2 \times \text{valhoogte} / 9,81)}$.

2.5 Data voorbereiding

De kwaliteit van de gegevens is belangrijk om zo betrouwbaar mogelijk eventuele relaties tussen de nagroei-indicatoren en parameters van de bedrijfsvoering vast te stellen. Daarom is veel inspanning geleverd om deze kwaliteit zo goed mogelijk te krijgen. Hiervoor is de kwaliteit van de aangeleverde informatie gecontroleerd voorafgaand aan de data analyse. De datavoorbereiding omvat met name het controleren van de aangeleverde informatie, het uitvoeren van verschillende consistentie checks (voor één locatie, en tussen verschillende locaties), het corrigeren van tyfouten en het verwijderen van tekst strings bij getallen, het combineren van de door de bedrijven aangeleverde informatie en de gegevens uit Rewab, het reorganiseren van alle afzonderlijke formats tot één overall database op zodanige wijze dat deze ingelezen kan worden in de analyse software (Python). Daartoe zijn unieke coderingen opgesteld voor alle eigenschappen. Deze kwaliteitscontrole en datavoorbereiding is deels automatisch uitgevoerd door te werken met een eenduidig format dat restricties kent op de invulvelden (bijvoorbeeld drop down menu's) en deels door de data handmatig te controleren.

De Excel database bevat de eigenschappen van 43 locaties. In totaal zijn er 1144 eigenschappen geïdentificeerd, verdeeld over de categorieën ontwerp, bedrijfsvoering en waterkwaliteit. De locaties kunnen beschreven worden aan de hand van een veel kleiner aantal eigenschappen, omdat een locatie bijvoorbeeld vaak maar twee beluchtingsstappen

bevat (bijvoorbeeld een cascade en een sproeier) waardoor de eigenschappen van de andere beluchtingstypes (plaatbeluchting, BOT beluchting, venturi) niet relevant zijn voor die locatie.

Per locatie zijn gemiddeld circa 400 eigenschappen bekend, het minimum is 225 en het maximum is 680. Ongeveer 15 – 30% van deze eigenschappen zijn waterkwaliteitsgegevens (gemiddeld 100 waterkwaliteitsparameters per locatie), de overige gegevens zijn ontwerp en bedrijfsvoeringsgerelateerde eigenschappen. In Python is het aantal eigenschappen van een locatie uitgebreid met een beperkt aantal afgeleide parameters (zogenaamde *lazy attributes*), zie paragraaf 2.4.3.

Het maximale aantal te vormen 1-op-1 relaties tussen 2 eigenschappen is voor elke locatie ongeveer 85.000 ($1/2 \times 410^2$). Om focus te kunnen aanbrengen is het aantal parameters dat in deze relatie analyse betrokken is op verschillende manieren gereduceerd:

- Een aantal parameters is op voorhand niet in de analyse betrokken omdat:
 - het aantal datapunten te beperkt is;
 - de waarden van de eigenschap niet of nauwelijks onderscheidend is voor de 43 productielocaties;
 - de informatie kwalitatief van aard is, en daarmee niet uit te drukken in een getal of een TRUE of FALSE statement. Deze informatie kan als aanvullende informatie nuttig zijn om een observatie nader te verklaren;
 - de keuze gemaakt is om een beperkt aantal gegevens nog niet in deze analyse te betrekken, deze informatie kan als aanvulling nuttig zijn om een observatie nader te verklaren.

In Bijlage I is voor het grootste deel van de parameters omschreven of en op welke manier ze in de analyse betrokken zijn.

- In de Excel database is voor de locaties die uit meerdere filter- of beluchtingsstappen in serie of in parallelle schakeling bestaan de eigenschappen voor alle afzonderlijke filterspannen en beluchtingstappen geaggregeerd tot één eigenschap per locatie. Daarmee wordt de omvang van de correlatie analyse sterk gereduceerd. Tevens is het hiermee mogelijk om een locatie die uit één beluchtingsstap en één filtratiestap bestaat te vergelijken met een locatie die twee parallelle straten heeft met elk meerdere beluchtings- en filtratiestappen in serie (zie paragraaf 2.5.1).
- In de Excel database zijn daarnaast verschillende waterkwaliteitsparameters uitgesloten:
 - In een enkel geval is sprake van meerdere ruwwatertakken, deze zijn verwerkt tot één ruwwaterkwaliteit per locatie.
 - De waarden van verschillende waterkwaliteitsparameters (ruw en rein) zijn niet onderscheidend, veelal omdat deze op of net boven de rapportagegrens zitten. Deze uitgangspunten zijn omschreven in paragraaf 2.4.1.1 en Bijlage III.

Na de reductie in de Excel database blijven er circa 130 eigenschappen over. In Python zijn daar verschillende afgeleide parameters aan toegevoegd. In totaal zijn in Python circa 23.000 relaties getest en de significantie is uitgedrukt aan de hand van de Spearman correlatie coëfficiënten (zie paragraaf 2.6).

2.5.1 Data aggregatie

Het format verzamelt de informatie per beluchtingsstap en per filterstap. Circa 30 van de 43 productielocaties bestaat uit één straat. De overige locaties bestaan uit parallelle straten die tijdens het proces of in de reinwaterkelder bij elkaar komen. Binnen een locatie is in veel gevallen sprake van meerdere beluchtingsstappen en/of filtratiestappen in serie. De

informatie die beschikbaar is op unit- en straatniveau is geaggregeerd tot locatieniveau om alle locaties goed met elkaar te kunnen vergelijken waardoor het aantal zuiveringen in de database gemaximaliseerd wordt. Als voorbeeld: door middel van deze aggregatie is het mogelijk om een zuivering met 1 straat die bestaat uit 1 beluchtingsstap en 1 filterstap te vergelijken met een zuivering die is opgebouwd uit 2 parallelle straten, die elk weer bestaan uit 2 beluchtingsstappen in serie (eventueel verschillend van type) en 2 filterstappen in serie.

De aggregatie is aan de hand van verschillende regels consistent uitgevoerd (zie ook Bijlage I). Enkele regels zijn:

- Contact- en verblijftijden van stappen in serie zijn gesommeerd. Contact- en verblijftijden van parallelle stappen zijn debietsproportioneel gemiddeld.
- Filtratiesnelheden zijn debietsproportioneel gemiddeld.
- Spoelwater- en spoellucht volumes zijn gesommeerd voor stappen in serie, en debietsproportioneel bepaald voor parallelle stappen. Voor de gehele locatie geldt de som van de afzonderlijke parallelle straten.
- Debieten zijn capaciteitsproportioneel gemiddeld.
- Lucht-water verhoudingen van stappen in serie zijn gesommeerd. Lucht-water verhoudingen van parallelle stappen zijn debietsproportioneel gemiddeld.
- Het effectief beluchtingsgetal van stappen in serie zijn gesommeerd. Het effectief beluchtingsgetal van parallelle stappen zijn debietsproportioneel gemiddeld.
- Parameters zoals de afstand van de spoelgoten tot aan het filterbed, het aantal nozzles per m², het aantal spoeldoppen per m², etc. zijn debietsproportioneel gewogen.
- De locatie waarde voor binaire parameters (TRUE/FALSE) is vaak gebaseerd op de meerderheid onder afzonderlijke stappen. Details over uitzonderingen staan in Bijlage I.

De volgende aggregaties hebben op locatieniveau plaatsgevonden:

- Fikkersdries I en Fikkersdries II zijn als twee locaties ingevuld. De reinwaters worden gemengd en daarna gedistribueerd. Daarom zijn deze twee locaties verwerkt tot één locatie "Fikkersdries I + 2".
- Lekkerkerk Tiendweg en Lekkerkerk Schuwacht zijn als twee locaties ingevuld. De reinwaters worden gemengd en daarna gedistribueerd. Daarom zijn deze twee locaties verwerkt tot één locatie "Lekkerkerk Totaal".
- Helmond heeft twee ruwwaterstrengen en twee daarop aangepaste zuiveringen. De ruwwater zijn debietsproportioneel gemiddeld.

2.6 Data-analyse

Binnen dit project is een data-analyse uitgevoerd met als doelstelling het analyseren van verzamelde data afkomstig van drinkwaterproductiebedrijven en het duiden van gevonden verbanden die mogelijk relevant zijn om de kennis over grondwaterzuiveringen te vergroten.

De onderzochte relaties worden in sommige gevallen verstoord door zogenaamde '*disturbances*' of verstoringen. Voorbeelden van verstoringen zijn de invloed van extra zuiveringsstappen naast de beluchting en filtratie, deelstroombehandeling, het bijmengen van reinwater van een andere productielocatie en het effect van het distributienet op de waterkwaliteit aldaar. Bekend is dat zulke verstoringen aanwezig zijn, ten dele is bekend om welke verstoringen het gaat. Het effect van deze verstoringen is in dit onderzoek niet meegenomen vanwege de onbekendheid van de verstoringen en de complexiteit van het onderzoek. Bij het zoeken naar relaties tussen waterkwaliteitsparameters en proceseigenschappen is het van belang om te realiseren dat de invloed van de eventuele aanwezigheid van dergelijke verstoringen de uitkomsten kunnen vertroebelen. Zo kan het

zijn dat (een) bepaalde zuiveringsparameter(s) gecorreleerd is/zijn met de nagroei-indicatoren, maar dat door (een) versturende factors deze relatie niet wordt gevonden.

2.6.1 Kenmerken dataset en uitgangspunten analyse

Voor de data analyse is een dataset met een grote variëteit aan gegevens beschikbaar, gevoed door gegevens afkomstig van 43 grondwaterzuiveringen (instanties). Er zijn ca. 1000 eigenschappen (attributen) van een zuivering gedefinieerd, zie paragraaf 2.4. De dataset wordt gekenmerkt door:

1. Heterogene variëteit in attributen: de onderzochte grondwaterzuiveringen hebben niet allemaal dezelfde zuiveringstappen, of er vindt binnen een productielocatie menging plaats van water met water die een andere behandeling heeft ondergaan. Hierdoor zijn gegevens (attributen) niet eenvoudig vergelijkbaar. Voorbeelden zijn de mogelijke invloed van extra zuiveringsstappen naast de beluchting en filtratie, deelstroombehandeling, het bijmengen van reinwater van een andere productielocatie en het effect van het distributienet op de waterkwaliteit benedenstreams.
2. Hiërarchische eigenschappen (overerving): veel eigenschappen kunnen slechts een waarde hebben als een voorgaande eigenschap een waarde heeft. Voorbeeld: alleen indien er sprake is van een cascadebeluchting zijn alle daaraan gerelateerde ontwerpparameters relevant in de analyse (valhoogte, aantal valstappen etc.).
3. Deels vanwege de hiërarchie en deels vanwege het ontbreken van gegevens bij een aantal zuiveringen zijn niet voor elke attribuut 43 gegevens beschikbaar. De volledige matrix aan attributen is daardoor 'dun' gevuld (de matrix is *sparse*).
4. Relatief veel bekende eigenschappen per zuivering in relatie tot het aantal zuiveringen (in data-analysetermen: *veel attributen per instantie*). Dit betekent dat er relatief veel variabelen zijn, een nog groter aantal (enkelvoudige) relaties mogelijk zijn, en slechts zeer weinig observaties beschikbaar die mogelijke verbanden statistisch solide kunnen onderbouwen.

Door de grote variëteit aan informatie en een klein aantal onafhankelijke observaties (instanties), wordt een data-analyse bemoeilijkt. Ook gezien de overige eerdergenoemde kenmerken (hiërarchie en dunne vulling) is het trainen van complexere, *machine-learning* regressiemodellen niet de aangewezen analyseroute. De volgende aanpak is daarom gekozen:

- berekening en visualisatie van de Spearman correlatiematrix. Voor deze (enkelvoudige) relaties is de Spearman's rank correlatiecoëfficiënt berekend. Deze correlatiecoëfficiënt geeft een maatstaf tussen +1 en -1, afhankelijk van hoe sterk de variabelen met elkaar correleren, gebaseerd op rangorde in waarden. Daarbij betekent een waarde van -1 een geheel negatieve correlatie, en +1 betekent een positieve correlatie. Het verschil met de Pearson correlatiecoëfficiënt is dat de data niet normaal verdeeld hoeft te zijn en dat zowel lineaire als niet-lineaire relaties worden aangetoond;
- het onderzoeken van hypothese gedreven selecties van attributen, om daarmee (na eventuele visualisatie) mogelijke aanknopingspunten te vinden voor vervolgonderzoek. Een deel van de hypothese casussen is door W. de Vet aangedragen;
- automatische clusteringsanalyse om datapunten met soortgelijke eigenschappen te groeperen.

Belangrijk in de voorgestelde aanpak is de inzet van domeinkennis om relevante selecties uit de dataset te maken en kennis van analysetechnieken. Bij het uitvoeren van de gegevensanalyse dient in het achterhoofd gehouden te worden dat er impliciete (cor)relaties kunnen zijn tussen ruwwaterkwaliteit, procesontwerp en bedrijfsvoering (multicollineariteit).

Voor de gevonden relaties is dit project geen onderzoek gedaan naar de aanwezigheid een causaal verband, noch zijn mogelijk achterliggende oorzaken bij de betreffende relatie onderzocht.

2.6.2 Databewerking

Aanvullend op de parameters die reeds in de aangelegde dataset verwerkt zijn, is tevens een set van afgeleide eigenschappen gedefinieerd. Middels domeinkennis zijn ontwerpparameters daarin gecombineerd tot parameters die een vergelijking tussen diverse zuiveringstappen mogelijk maakt. Met deze afgeleide eigenschappen wordt de hiërarchie van de dataset deels ingeperkt (omdat verschillende zuiveringsstappen nu onderling vergeleken kunnen worden) en worden meerdere parameters samengevoegd tot een meer generieke eigenschap. Deze afgeleide eigenschappen zijn in paragraaf 2.4.3 beschreven. De afgeleide eigenschappen vormen samen met de oorspronkelijke dataset het startpunt voor de verdere analyse.

Binnen de gekozen hypothese gedreven benadering is gebruik gemaakt van een *machine-learning* toolbox voor het verwerken van de data. KWR heeft de toolbox uitgewerkt in de programmeertaal Python, waarbij gebruik wordt gemaakt van bestaande opensource pakketten pandas, bokeh en Scikit-Learn (Pedregosa et al, 2011). De toolbox heeft meerdere functies: inladen van data uit bronbestand(en), afgeleide eigenschappen berekenen en toevoegen, en voorbewerking (o.a. instanties of attributen met incomplete informatie verwijderen). Ook biedt de toolbox functies met betrekking tot statistiek, regressie-analyse en visualisatie van data.

2.6.3 Clustering

Hoewel de dataset niet geschikt is om regressiemodellen te trainen en valideren, zijn selecties van de dataset wel gebruikt om clustering op uit te voeren. Clustering is een *machine-learning* techniek waarbij instanties in een dataset worden gegroepeerd in clusters, op basis van een afstandscriterium. Hoe groter de afstand tussen van een punt tot een andere groep (meet)punten, hoe groter de kans dat dit specifieke punt tot een apart cluster wordt gerekend. Clustering biedt op die manier de mogelijkheid om zuiveringen met overeenkomstige eigenschappen in een apart cluster in te delen, terwijl andere zuiveringen die andere effecten veroorzaken, zullen worden gegroepeerd in een ander cluster. Een clusteranalyse kan interessant zijn om snel een indruk te krijgen van de mate van (effect)verschil tussen locaties in de dataset. Een dergelijke analyse kan dan ook het startpunt zijn om zuiveringen in een specifiek cluster verder te analyseren. Voor de clusteranalyses is gebruik gemaakt van het zogenaamde k-Means algoritme en Gaussian Mixture Models (GMM). Een meer uitgebreide omschrijving van clustering als techniek, en de gebruikte algoritmes is te vinden in (Vonk, E. et al, 2015).

2.6.4 Visualisatie

Als verdere verkenning van de dataset zijn vier typen visualisaties gemaakt, deels op basis van vooraf gedefinieerde hypothesen:

- **4D-scatterplots** Een visualisatie waarbij een variabele op de horizontale grafiekas wordt uitgezet tegen een variabele op de verticale grafiekas. Om daarnaast twee andere eigenschappen in deze visualisaties te vangen is ook gebruik gemaakt van kleuring van de datapunten en de grootte van datapunten. Bij continue variabelen is telkens de laagste waarde groen gekleurd en de hoogste waarde in de set met rood (andere kleuren zoals oranje en geel vallen dan in het kleurspectrum daartussen). De grootte van een datapunt correspondeert met de grootte van een waarde.

- Scatterplot matrix Voor selecties op de dataset zijn alle parameters stuksgewijs tegen elkaar uitgezet waardoor een verzameling aan scatterplots ontstaat. Hiermee kunnen enkelvoudige relaties gevisualiseerd worden (1-op-1 relaties). Relaties met een hoge Spearman correlatiecoëfficiënt (groter dan ca. 0,5 of kleiner dan ca. - 0,5) en die praktische relevantie hebben zijn in Python of in Excel gemaakt.
- Correlation heatmap Een zogenaamde *correlation heatmap* is gebaseerd op hetzelfde principe als de scatterplot matrix (het stuksgewijs tegen elkaar uitzetten van parameters), met als verschil dat hierbij niet de datapunten zelf worden weergegeven, maar alleen de waarden van de correlatiecoëfficiënten. Deze coëfficiënten zijn middels een oplopende kleurschaal gevisualiseerd. Een dergelijke visualisatie biedt de mogelijkheid om snel de meest sterke correlaties uit te zoeken. Omdat in deze database echter (na reductie) circa 10.000 relaties beschikbaar zijn is ook een lijst gebruikt waarin de correlatie coëfficiënten of grootte gesorteerd zijn. Op basis van de *heatmap* en deze lijst is een aantal scatterplots gemaakt (interessant in het licht van dit onderzoek en een redelijk hoge correlatie coëfficiënt).
- 3D-scatterplots. Voor een aantal casussen is hypothese-gedreven analyse verricht van een combinatie van twee mogelijk verklarende factoren. Daartoe zijn verschillende waterkwaliteitsparameters (bijvoorbeeld het KG22 in het gedistribueerde water) geplot tegen een andere (continue) variabele die voor een groot aantal locaties bekend is (bijvoorbeeld het effectief beluchtingsgetal), en is in die visualisatie een derde informatielaag toegevoegd met parameters die een binaire (logische) waarde hebben (TRUE/FALSE, zoals wel of geen drukfiltratie). Om een mogelijk interessante binaire parameter te kiezen is de volgende aanpak gevolgd:
 1. Reduceer de database tot de locatie-geaggregeerde binaire eigenschappen (TRUE/FALSE waarde) en verschillende waterkwaliteit gerelateerde parameters.
 2. Orden vervolgens de data per kwaliteitsparameter van klein naar groot. Beoordeel op kwalitatieve wijze of er een verband lijkt te zijn tussen de oplopende waarde van de waterkwaliteit parameter en het voorkomen van de TRUE of FALSE waarden.
 3. In geval dat een verband lijkt te bestaan zijn de betreffende datapunten gevisualiseerd. Deze methode is zeer tijdrovend vanwege de grote keuze van parameters en is slechts voor enkele kwaliteitsparameters getest, namelijk voor KG22 en *Aeromonas* in het gedistribueerde water, en het ijzergehalte, het TOC gehalte en de troebelingsgraad in reinwater.

2.6.5 Correlatie-analyse

De correlaties en aanwijzingen op verbanden die middels de huidige aanpak zijn gevonden sluiten niet uit dat er daarnaast nog meer (cor)relaties bestaan. Het is zeer goed mogelijk dat twee of meer parameters afzonderlijk geen enkele correlatie vertonen met een derde parameter, maar dat deze parameters tezamen wel een correlatie hebben. Dit staat bekend als de 'XOR'-relatie (*two class exclusive* relatie, Domingos, 2012)). Om XOR-relaties aan te tonen is een multiregressie analyse nodig. Bij een dergelijke analyse komen er extra vrijheidsgraden die gevalideerd moeten worden, omdat elke extra verklarende parameter een nieuwe vrijheidsgraad introduceert. Dit kenmerk vereist dat er een toenemend aantal onafhankelijk aantal datapunten (hier: grondwaterzuiveringen) beschikbaar is om deze correlatie te bewijzen. Om die reden is besloten multiregressie analyses met XOR-relaties niet uit te voeren in dit onderzoek. In SPSS kan een dergelijke analyse worden uitgevoerd, maar dat moet grotendeels handmatig plaatsvinden. Dat vergt daardoor veel tijd, om die reden is besloten om deze analyse (nog) niet uit te voeren.

3 Resultaten

3.1 Inleiding

Het onderzoek heeft verschillende (deel)resultaten opgeleverd. Allereerst is een rijke en waardevolle database opgeleverd die gebaseerd is op i) de inventarisatie van toegepaste beluchting- en filtratieprocessen (de procesconfiguratie, het procesontwerp en de bedrijfsvoeringscondities) van 43 Nederlandse grondwaterpompstations die anoxisch grondwater behandelen, en ii) de ruwwaterkwaliteit, de reinwaterkwaliteit en de waterkwaliteit van het gedistribueerde water. Deze database is in voorliggend onderzoek gebruikt om naar correlaties te zoeken tussen nagroei indicatoren en relevante fysisch-chemische parameters van het reinwater en het gedistribueerde water enerzijds en de ontwerp en bedrijfsvoeringsgerelateerde eigenschappen anderzijds. De database kan van grote waarde zijn bij vervolgonderzoek naar biologische activiteit of bij andere drinkwaterbereidingsgerelateerde onderzoeken, bijvoorbeeld als naslagwerk, om de prestatie en eigenschappen van zuiveringen te vergelijken, en bij onderzoek, ontwerp en optimalisatie van grondwaterzuiveringen in bredere zin.⁶ In de navolgende paragrafen wordt ingegaan op de verkregen inzichten in de effecten van klassieke zuiveringsprocessen in de grondwaterzuivering op de effectiviteit van het beperken van de nagroei indicatoren (KG22 en *Aeromonas*) en de verwijdering van typische grondwatergerelateerde fysisch-chemische componenten. Achtereenvolgens komen aan bod:

- Vergelijking van de significante correlaties nagroei indicatoren uit deze studie met die uit het BTO onderzoek van 2011;
- Relaties (enkelvoudig) nagroei indicatoren en chemische waterkwaliteit
- Relaties (enkelvoudig) tussen nagroei indicatoren en ontwerp zuivering
- Relaties (enkelvoudig) tussen nagroei indicatoren en bedrijfsvoering zuivering
- Relaties (meervoudig) nagroei indicatoren a.h.v. verschillende hypothesen
- Relaties (meervoudig) waterkwaliteit, continue variabele en binaire variabele
- Relaties (enkelvoudig) chemische waterkwaliteit en bedrijfsvoering

3.2 Relaties met chemische waterkwaliteit – vergelijk BTO 2011 onderzoek

In Tabel 7 en Tabel 8 zijn de correlatie coëfficiënten van de significante enkelvoudige relaties tussen chemische waterkwaliteit en nagroei indicatoren (KG22 en *Aeromonas*) uit van der Wielen (2011) vergeleken met de correlatie coëfficiënten in voorliggend onderzoek.

⁶ In een BTO onderzoeksvoorstel over ontijzering van de Themagroep 'Drinkwatertechnologie van de Toekomst' is voorgesteld om de database te gebruiken. Het voorstel is door het CO goedgekeurd en zal naar verwachting in 2016 starten.

TABEL 7. VERGELIJKING CORRELATIES CHEMISCHE PARAMETERS VERSUS KG22 GEDISTRIBUEERD WATER VAN DER WIELEN (2011) EN VOORLIGGEND ONDERZOEK

Parameter	Van der Wielen, 2011	Voorliggend onderzoek	
	R ² (lineaire regressie)	R ² (lineaire regressie)	Spearman
TOC – ruw	0,59	0,25	0,54
TOC – distributie	0,47	0,51	0,68
TOC – rein	0,47	0,24	0,38
Methaan – ruw	0,41	0,54	0,57
Ijzer – ruw	0,36	0,09	0,61
Ammonium – ruw	0,34	0,0006 (richting: -)	0,17

Uit deze vergelijking blijkt dat de relaties tussen het KG22 en het TOC gehalte van het reinwater en het ammonium gehalte van het ruwwater minder significant zijn voor de set van productielocaties die in dit onderzoek gebruikt zijn, maar de andere correlaties zijn vergelijkbaar tussen beide studies.

TABEL 8. VERGELIJKING CORRELATIES CHEMISCHE PARAMETERS VERSUS *AEROMONAS* GEDISTRIBUEERD WATER (VAN DER WIELEN, 2011) EN VOORLIGGEND ONDERZOEK

Parameter	Van der Wielen, 2011	Voorliggend onderzoek	
	R ²	R ² (lineaire regressie)	Spearman
TOC – ruw	0,45	0,039	0,29
Ammonium – ruw	0,44	0,01 (richting: -)	-0,02
TOC – distributie	0,42	0,15	0,37
TOC – rein	0,38	0,05	0,09
Methaan – ruw	0,19	0,13	0,34

Uit deze vergelijking blijkt dat de relaties tussen *Aeromonas* en het TOC gehalte van het reinwater en het ammonium gehalte van het ruwwater minder significant zijn voor de set van productielocaties die in dit onderzoek gebruikt zijn. De overige correlaties zijn vergelijkbaar tussen beide studies. De *Aeromonas* aantallen in Van der Wielen (2011) zijn loggetransformeerd, bij voorliggende studie is dat niet het geval.

De correlaties tussen het TOC gehalte in het gedistribueerde water en het methaan gehalte in ruwwater enerzijds en de groei indicatoren anderzijds komen in beide onderzoeken goed naar voren. Hieruit blijkt dat de significantie van de correlatie ondanks een redelijke groepsgrootte (respectievelijk 38 en 43) sterk afhangt van de set van productielocaties die in de regressieanalyse betrokken zijn, of van de periode waaruit de data afkomstig zijn.

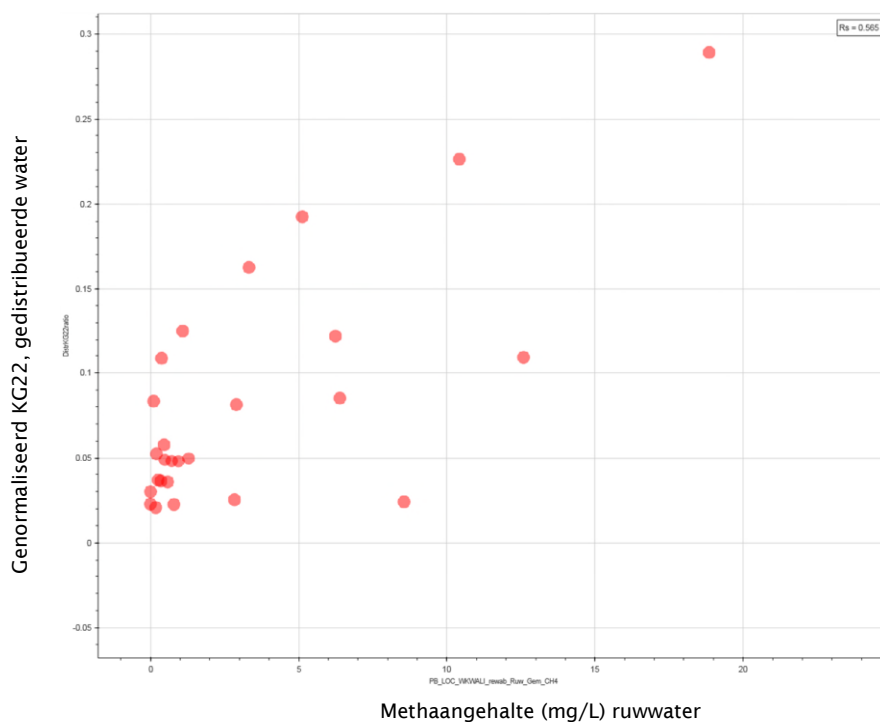
Deze relaties komen in de navolgende paragrafen terug in verschillende visualisaties (voor de dataset die in deze studie gebruikt is).

In voorliggende studie zijn ook de relaties tussen het ontwerp, de bedrijfsvoering en de chemische waterkwaliteit onderzocht omdat verondersteld is dat de aanwezigheid van zulke opgeloste en deeltjesgebonden stoffen een negatieve invloed kunnen hebben op de biologische stabiliteit van reinwater en bijdragen aan groei in het leidingnet. Met de toegepaste dataset zijn geen negatieve relaties gevonden tussen de waarden van de groei indicatoren in reinwater of gedistribueerd water en de waarden van het ijzergehalte of de troebelingsgraad in reinwater of gedistribueerd water.

3.3 Relaties nagroei indicatoren en chemische waterkwaliteit

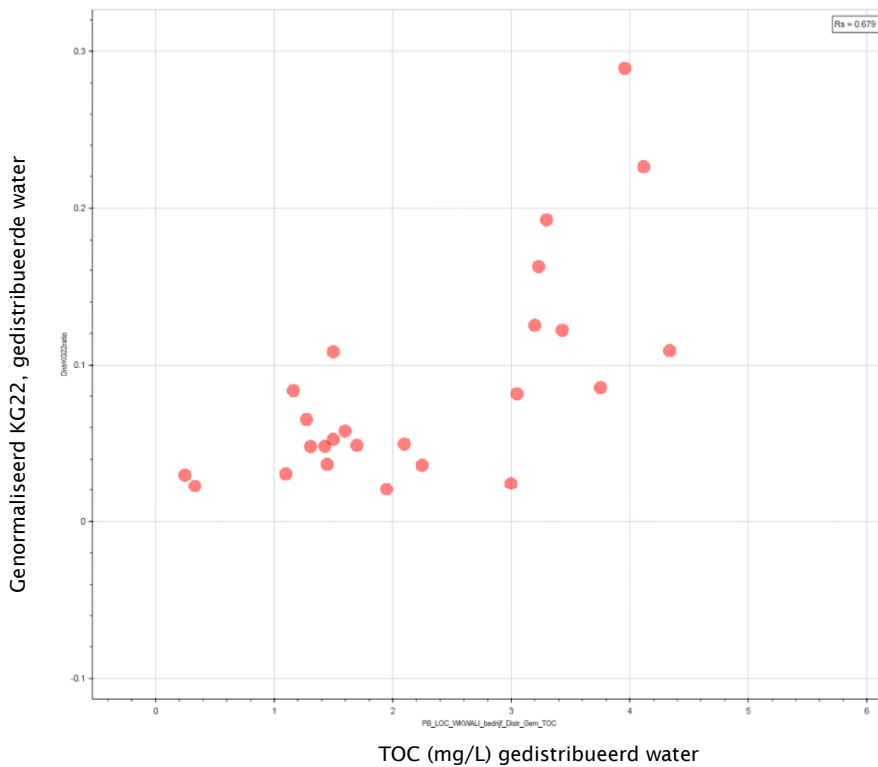
In deze paragraaf zijn verschillende relaties tussen de waarden van de nagroei indicatoren en de chemische waterkwaliteit besproken. De getoonde verbanden zijn geselecteerd op basis van relevantie, de waarde van de correlatie coëfficiënt en de relatie met voorgaand BTO onderzoek. Het volledige overzicht van de ca. 11.000 correlatie coëfficiënten is beschikbaar, maar niet in dit rapport opgenomen.

Figuur 2 toont een positief verband tussen het genormaliseerd KG22 (normalisatie t.o.v. normwaarde) van het gedistribueerde water en het methaan gehalte in ruwwater. De Spearman coëfficiënt van deze relatie is 0,57. Dit verband is ook in paragraaf 3.2 besproken.



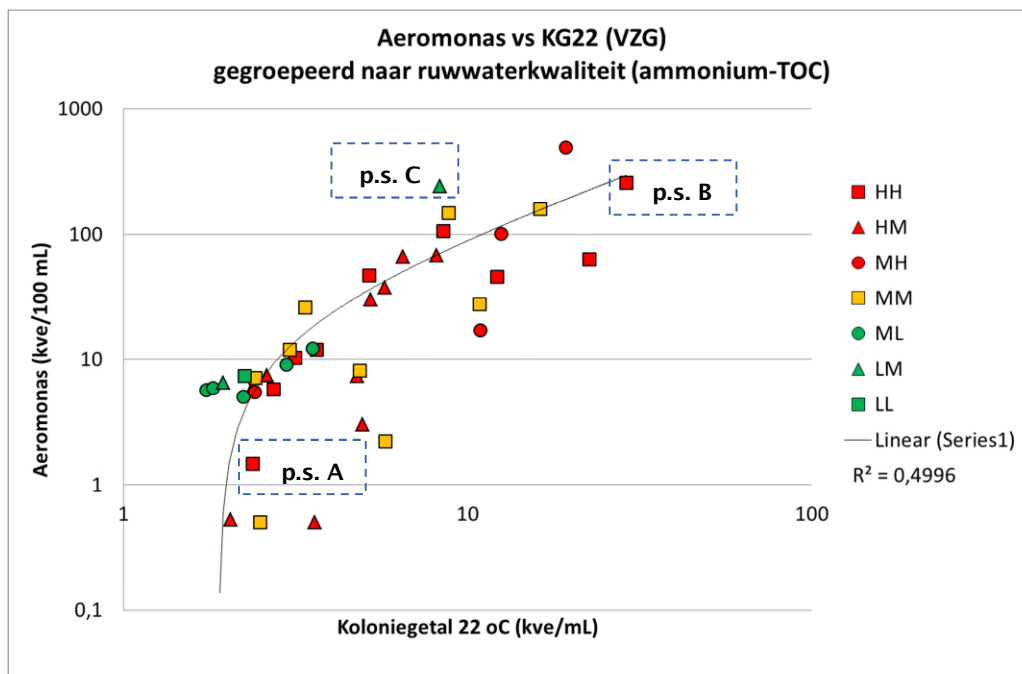
FIGUUR 2; GENORMALISEERD KG22 GEDISTRIBUEERD WATER VS METHAANGEHALTE RUWWATER.

Figuur 3 toont een positief verband tussen het genormaliseerd KG22 en het TOC gehalte van het gedistribueerde water. De Spearman coëfficiënt van deze relatie is 0,68. De relatie met het TOC in het ruwwater is redelijk sterk (Spearman coëfficiënt 0,54), echter de relatie met het TOC gehalte in het reinwater is zwak. Deze verbanden zijn ook in paragraaf 3.2 besproken.



FIGUUR 3; GENORMALISEERD KG22 GEDISTRIBUEERD WATER VS TOC GEHALTE GEDISTRIBUEERD WATER.

Figuur 4 toont de relatie tussen het KG22 en *Aeromonas* in het gedistribueerde water en de ruwwater classificatie van de productielocatie op basis van ammonium en TOC (conform Tabel 1). Dit is een alternatieve wijze om de verbanden die in paragraaf 3.2 zijn besproken te visualiseren.



FIGUUR 4. RELATIE (LOG-LOG PLOT) TUSSEN KG22 EN AEROMAS IN HET GEDISTRIBUEERDE WATER, EN RUWWATER CLASSIFICATIE A.H.V. AMMONIUM-TOC SCORE (ZIE ONDERSTAANDE TABEL VOOR CODERING). "P.S." STAAT VOOR POMPSTATION, ZIE REFERENTIE IN DE TEKST HIERNA.

Classificatie	Ammonium [mg/L]	Organisch koolstof (TOC of DOC) [mg/L]
L (laag)	< 0,1	< 1,0
M (midden)	0,1 – 0,5	1,0 – 4,0
H (hoog)	> 0,5	> 4,0

Uit Figuur 4 blijkt geen eenduidige relatie tussen de ruwwater classificatie en de nagroei indicatoren, maar op hoofdlijnen worden de in paragrafen 3.2 en 3.3 besproken relaties bevestigd. Zo lijken de productielocaties met een lager ammonium en/of TOC gehalte in het ruwwater (de groen gekleurde punten) lagere nagroei indicator waarden te hebben. Echter, er zijn ook locaties met een hoger ammonium en/of TOC gehalte in het ruwwater die toch lagere waarden voor de nagroei indicatoren hebben. De hoogste waarden van de nagroei indicatoren zijn aanwezig bij de productielocaties met een hoger ammonium en TOC gehalte in het ruwwater. Er lijkt een beperkt onderscheid te zijn tussen de groep met medium ammonium en TOC gehalten in het ruwwater (MM) en de groepen met één of twee hogere gehalten (HM, MH of HH). In paragraaf 3.4 is aan deze analyse verbreed door het type beluchting toe te voegen.

In Figuur 4 zijn als voorbeeld twee pompstations gemarkeerd (p.s. A en p.s. B) die een vergelijkbare ruwwaterclassificatie hebben maar die zeer verschillende *Aeromonas* waarden hebben. Ook heeft Pompstation C opvallend hoge *Aeromonas* waarden gelet op de ruwwaterclassificatie. Op basis van deze figuur kunnen de eigenschappen (ontwerp en

bedrijfsvoering) van dergelijke opvallende zuiveringen handmatig vergeleken worden om te achterhalen of er verschillende eigenschappen bestaan die mogelijk verklarend zijn voor deze opvallende waarden.

3.4 Relaties tussen nagroei indicatoren en ontwerp zuivering

In deze paragraaf zijn verschillende relaties tussen de waarden van de nagroei indicatoren en ontwerpparameters besproken. De getoonde verbanden zijn geselecteerd op basis van praktische relevantie en de waarde van de correlatie coëfficiënt. Een volledig overzicht van alle correlatie coëfficiënten is beschikbaar, maar niet weergegeven in deze rapportage.

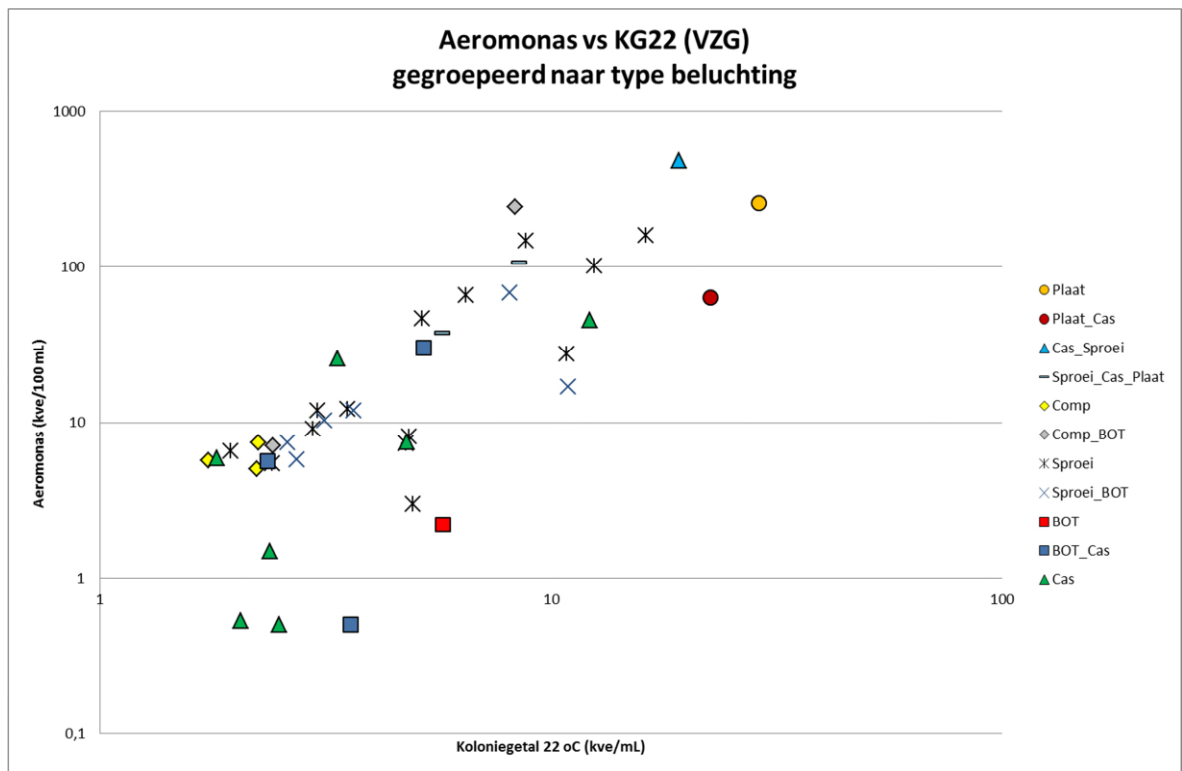
3.4.1 Type beluchting

Uit het overzicht van Tabel 9 blijkt dat er in de database sprake is van een brede spreiding van (combinaties van) typen beluchtingen van de productielocaties. De tabel geeft het productielocaties in de database weer aan de hand van type beluchting. Op de diagonaal van de tabel de locaties die één beluchtingsstap hebben. In deze paragraaf is beschreven of er een verband bestaat tussen het type beluchting en de waarde van de nagroei parameters.

TABEL 9. AANTAL LOCATIES MET BEPAALDE COMBINATIE VAN TYPEN BELUCHTINGSSTAPPEN

plaat			1		1
sproeier		6	1 + 1x +plaat	15	
cascade		3	7		
BOT	2	1			
compressor	4				
	compressor	BOT	cascade	sproeier	plaat

De aanwezige (combinaties) van beluchtingstypes zijn verwerkt in een plot met het KG22 en *Aeromonas* in het gedistribueerde water (Figuur 5).



FIGUUR 5. KG22 EN AEROMONAS IN HET GEDISTRIBUEERDE WATER EN TYPE BELUCHTING (LOG-LOG PLOT)

In Figuur 5 lijken de productielocaties met een plaatbeluchting (indicator in figuur: \circ en $-$) hogere nagroei indicator waarden te tonen, terwijl de productielocaties met cascade, BOT of compressor beluchting (indicator in figuur: Δ , \square , en \diamond) vaker lagere nagroei indicator waarden tonen. De productielocaties met sproeibeluchting (indicator in figuur: \times en $*$) bevinden zich in het gehele bereik, al tonen deze productielocaties relatief vaak hogere waarden voor de nagroei indicatoren. Deze observatie kan niet zondermeer in een duidelijke conclusie verwerkt worden omdat bij de ontwerpkeuze van het type beluchting kennis van de grondstof betrokken is. Zo hebben productielocaties met een complexere ruwwater samenstelling (hoger ammonium, TOC, methaan) naar verwachting een intensievere beluchting zoals plaatbeluchting. Het valt op dat de locaties met plaatbeluchting ondanks de intensiteit van deze beluchting toch hogere nagroei indicator waarden hebben. In de voorgaande paragrafen is de positieve relatie getoond tussen deze componenten in ruwwater en de waarde van nagroei indicatoren. Daartoe is hierna de connectie gemaakt tussen het type beluchting en de ruwwater classificatie op basis van ammonium en TOC conform Tabel 1:

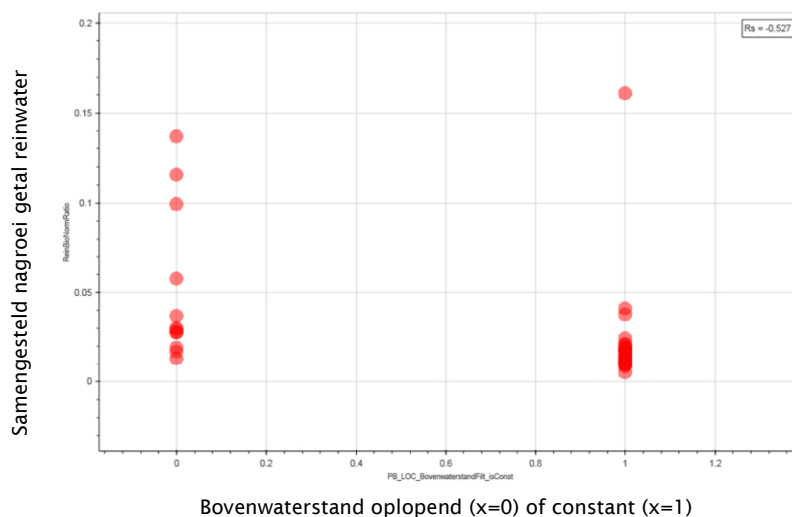
- Compressor beluchting komt in de lagere klassen voor (t/m MM).
- De cascade komt in alle hogere klassen voor, er is geen cascade bij de LL en de LM klasse.
- De BOT beluchting komt in alle hogere klassen voor, er is geen BOT beluchting in de LL en ML klassen.
- De plaatbeluchting komt alleen bij de hogere klassen voor (HH en HM).
- De sproeier komt in elke klasse behalve LL voor.

Dit overzicht geeft enig inzicht in hoe bij de ontwerp keuze van de beluchting rekening is gehouden met de ruwwatersamenstelling: in algemene zin is de intensiteit van de beluchting zwaarder naarmate het ruwwater meer ammonium en/of TOC bevat. De sproeier heeft een

beperkte effectiviteit en vormt daarmee een uitzondering omdat dit type beluchting ook bij hogere klassen wordt toegepast. Als de data vervolgens gerangschikt worden naar type beluchting en ruwwaterclassificatie en het *Aeromonas* aantal van het gedistribueerde water, en de *Aeromonas* aantallen binnen groepen met een gelijke ruwwater classificatie worden vergeleken, dan valt op dat de productielocaties met sproeiers relatief vaak hogere *Aeromonas* aantallen hebben. Een vergelijkbaar beeld geldt voor het KG22 van het gedistribueerde water. Dit is een aanwijzing dat sproeibeluchting minder gunstig is om biologisch stabiel water te kunnen produceren. Ook is er een aanwijzing dat BOT beluchting gunstiger lijkt dan plaatbeluchting voor de productie van biologisch stabiel water. Gerichte metingen naar methanotrofen (met qPCR) geeft duidelijkheid of biomassa van methanotrofen hierbij een rol speelt.

3.4.2 Bovenwaterstand constant of oplopend

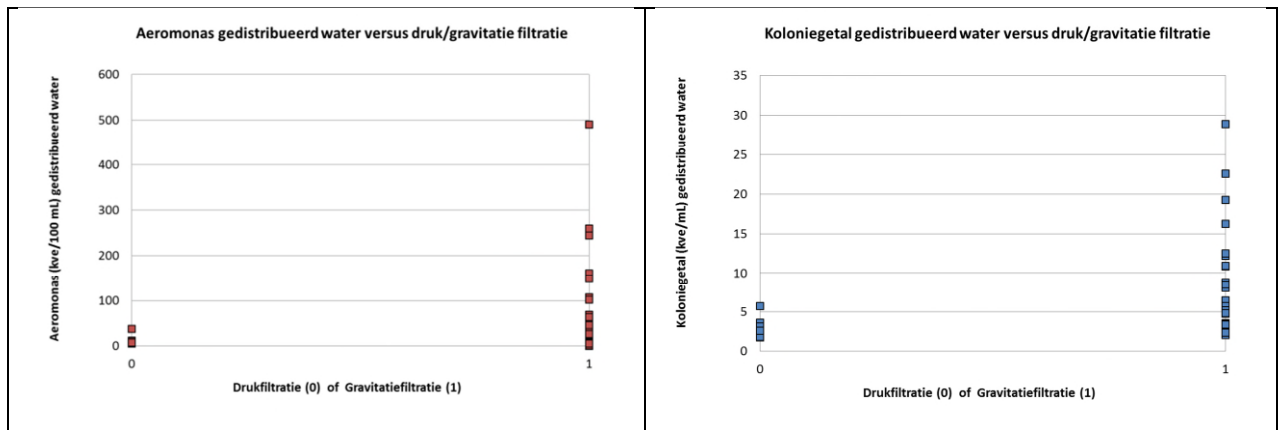
In Figuur 6 is het genormaliseerd samengesteld nagroei getal van het reinwater geplot tegen de situatie t.a.v. de bovenwaterstand: oplopende bovenwaterstand ($x=0$) of constante bovenwaterstand ($x=1$). De Spearman coëfficiënt is $-0,53$. De figuur laat zien dat de kans op hogere waarden van het nagroei getal groter is wanneer een productielocatie een oplopende bovenwaterstand heeft. Dit is een aanwijzing dat een constante bovenwaterstand gunstiger is om biologisch stabiel water te produceren.



FIGUUR 6. GENORMALISEERD SAMENGESTELD NAGROEI GETAL VOOR REINWATER VS OPLOPENDE BOVENWATERSTAND ($X=0$) OF BOVENWATERSTAND CONSTANT ($X=1$).

3.4.3 Drukfiltratie of gravitatie filtratie

De database bevat 12 locaties met drukfiltratie en 33 locaties met gravitatie filtratie. De productielocaties met drukfiltratie hebben allen relatief lage *Aeromonas* aantallen en een laag KG22 in het gedistribueerde water. Dit verband is getoond in Figuur 7. De Spearman coëfficiënten voor deze relaties zijn respectievelijk 0,29 en 0,55. Een vergelijkbaar beeld geldt voor de relatie met het KG22 in reinwater, echter voor *Aeromonas* in reinwater is het verband niet gevonden. Dit is een aanwijzing dat drukfiltratie gunstiger is voor de productie van biologisch stabiel water dan gravitatiefiltratie.



FIGUUR 7. *AEROMONAS* (LINKS) EN KG22 (RECHTS) GEDISTRIBUEERD WATER VERSUS DRUKFILTRATIE (0) OF GRAVITATIEFILTRATIE (1).

3.4.4 Resterend ontwerpparameters

De waarde van Spearman coëfficiënt voor de relatie tussen het KG22 van het gedistribueerde water en de leeftijd van de productielocaties na grootschalige renovatie of aanpassing is redelijk (0,64). Echter, bij slechts 17 van de 43 productielocaties is sprake van een dergelijke “tweede” leeftijd, en wanneer de leeftijd van de locaties wordt toegevoegd die niet zijn gewijzigd (de “oorspronkelijke” leeftijd) wordt de correlatie zeer zwak. Daarmee is de correlatie met de leeftijd na grootschalige renovatie of aanpassing niet relevant.

Alle relaties tussen de nagroei indicatoren in reinwater of het gedistribueerde water en ontwerpgerelateerde eigenschappen die een redelijk sterke correlatie coëfficiënt tonen én die een praktisch belang hebben, zijn hierboven benoemd. Er zijn geen sterke correlaties gevonden tussen nagroei indicatoren en de andere ontwerpgerelateerde eigenschappen. Een samenvattend overzicht van de ontwerpgerelateerde parameters is weergegeven in Tabel 6 en Bijlage I toont een completer overzicht. Uit die overzichten kan worden afgeleid met welke ontwerpgerelateerde eigenschappen in voorliggende studie geen relaties aangetoond zijn met de nagroei parameters. Uit dat volledige overzicht zijn hierna verschillende ontwerpeigenschappen geselecteerd waarvan op voorhand de verwachting bestond dat er mogelijk een correlatie bestaat met de nagroei indicatoren. In deze studie is geen verband gevonden tussen de volgende ontwerpgerelateerde parameters en nagroei indicatoren:

- De afstand van het filterbed tot aan de spoelgoten. Een grote afstand tot aan de spoelgoot kan mogelijk leiden tot een ineffektieve of onvolledige afvoer van deeltjes uit het filterbed, wat aanleiding zou kunnen geven tot nagroei.
- De aanwezigheid van luchtbehandeling. Hierbij is sprake van weinig onderscheidende data. Bij één van de 43 locaties wordt de lucht niet behandeld, en in drie gevallen is sprake van minimale behandeling met vliegengaas. De betreffende productielocaties hebben geen opvallend hoge nagroei parameters.
- De aanwezigheid van een ruwwater kelder. De aanwezigheid van een ruwwaterkelder zou negatief kunnen uitpakken op de waterkwaliteit door vroegtijdige oxidatie van ijzer en een minder effectieve werking van het filtratieproces. Het aantal productielocaties in de dataset met een ruwwater kelder is zeer gering.

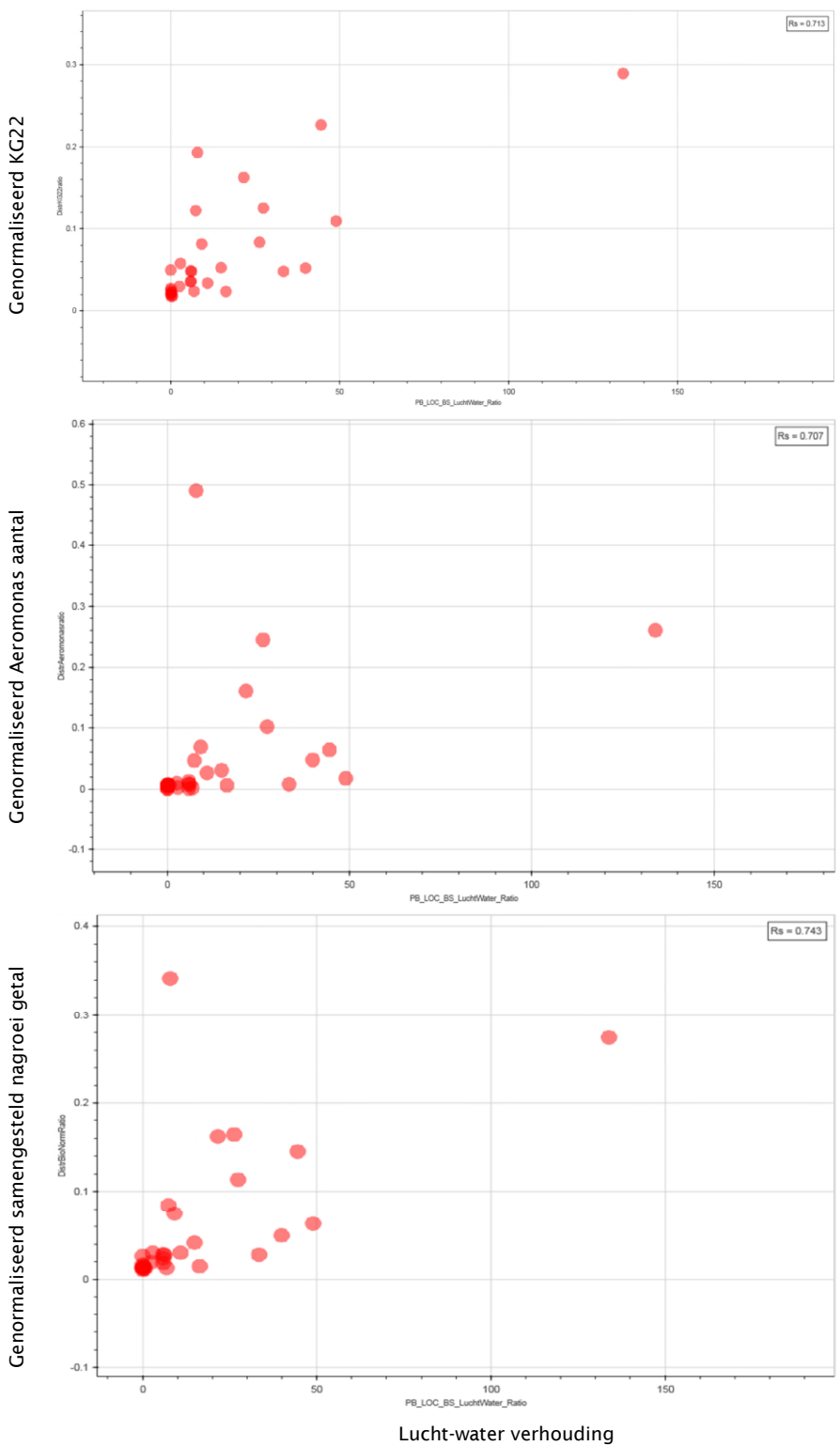
De afwezigheid van (één op één) relaties hoeft niet te betekenen dat parameters definitief geen invloed hebben op nagroei-indicatoren. De relatie kan door verstoringen door andere parameter(s) zijn weggevallen of is in combinatie met andere parameters (meervoudige regressie) wel van belang.

3.5 Relaties tussen nagroei indicatoren en bedrijfsvoering zuivering

In deze paragraaf worden verschillende relaties tussen de waarden van de nagroei indicatoren en bedrijfsvoeringsgerelateerde eigenschappen besproken. De getoonde verbanden zijn geselecteerd op basis van praktische relevantie en de waarde van de correlatie coëfficiënten. Een volledig overzicht van de ca. 11.000 correlatie coëfficiënten is beschikbaar maar wordt in dit rapport niet gepresenteerd.

3.5.1 Lucht-water verhouding

Figuur 8 toont de positieve verbanden tussen het genormaliseerd KG22, *Aeromonas*, en het samengesteld nagroei getal van het gedistribueerde water en de lucht-water verhouding. De Spearman coëfficiënten zijn respectievelijk 0,71, 0,71 en 0,74.



FIGUUR 8. GENORMALISEERD KG22, *AEROMONAS* EN SAMENGESTELD NAGROEI GETAL VOOR GEDISTRIBUEERD WATER VS LUCHT-WATER VERHOUDING.

Deze figuren tonen een positief verband tussen de waarde van de nagroei indicator (gedistribueerde water) en de lucht-water verhouding van de beluchting (geaggregeerd op locatie niveau). Dit is een aanwijzing dat een lagere lucht-water verhouding een positief

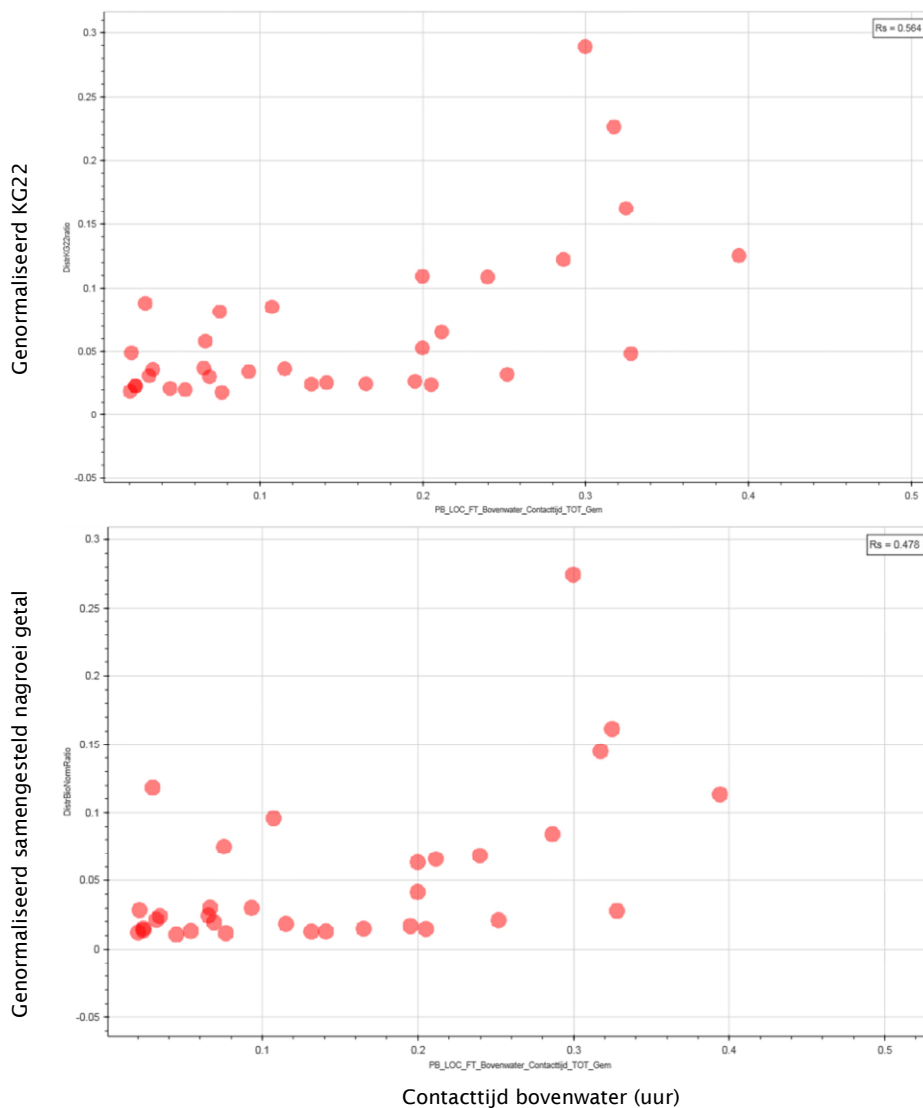
effect heeft op het beperken van de nagroei-indicatoren. Deze bevinding kan met twee mogelijke hypothesen uitgelegd worden: 1) vanwege de ruwwater samenstelling is een hoge lucht-water verhouding noodzakelijk, de verwijdering van chemische componenten (zoals ijzer, mangaan en ammonium) verloopt daardoor verondersteld effectief, wat leidt tot meer biomassa vorming in het filter en daardoor hogere waarden van nagroei in het gedistribueerde water, of 2) een hogere lucht-water verhouding is (mede) verantwoordelijk voor een hogere nagroei doordat bijvoorbeeld AOC gevormd wordt in de beluchting of doordat componenten (anders dan zuurstof) die in de lucht aanwezig zijn ondanks luchtbehandeling aan het water worden overgedragen en aldaar tot nagroei kunnen leiden.

3.5.2 Contacttijd bovenwater

Figuur 9 toont de positieve verbanden tussen het genormaliseerd KG22 en het samengesteld nagroei getal van het gedistribueerde water en de contacttijden van het bovenwater van de filters. De Spearman coëfficiënten zijn respectievelijk 0,56 en 0,48.

Deze observatie vormt een aanwijzing dat een lage bovenwater contacttijd mogelijk gunstiger is voor het verlagen van KG22 dan een hogere contacttijd. Eenzelfde richting geldt ook voor *Aeromonas*, maar die relatie is minder sterk. Mogelijk is er in geval van sproeibeluchting een verband met de contacttijd van water en lucht: deze is lager bij een hogere bovenwaterstand en een hogere bovenwater contacttijd. Een andere hypothese is dat er in het bovenwater processen optreden die een nadelig effect hebben op de biologische stabiliteit van het water.

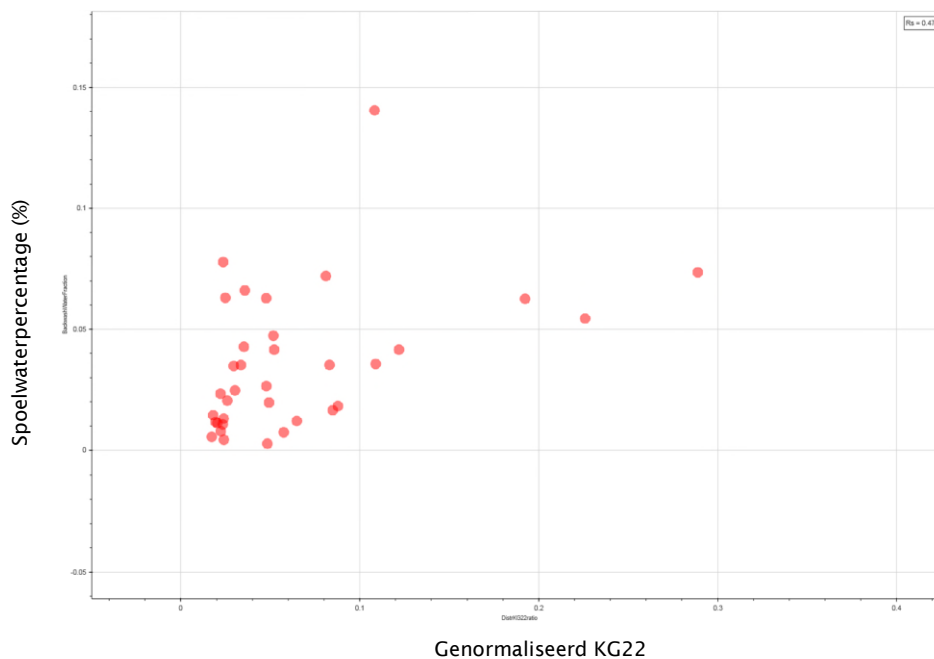
Overigens is ook een positief verband gevonden tussen de bovenwater contacttijd en de verwijdering van ijzer. Daarbij is niet onderzocht van welk type ontijzering sprake is (homogeen, heterogeen of biologisch). De betere verwijdering van ijzer kan ook de reden zijn van de lagere KG22-aantallen, omdat positief geladen ijzer in staat is negatief geladen organisch materiaal te immobiliseren in het distributiesysteem. Het SPO onderzoek van PWN en Evides heeft laten zien dat ijzer daarbij de biomassa-accumulatiesnelheid van het water kan verhogen (van der Kooij & Bereschenko, 2015).



FIGUUR 9. GENORMALISEERD KG22 EN SAMENGESTELD NAGROEI GETAL VOOR GEDISTRIBUEERD WATER VS CONTACTTIJD BOVENWATER (UUR).

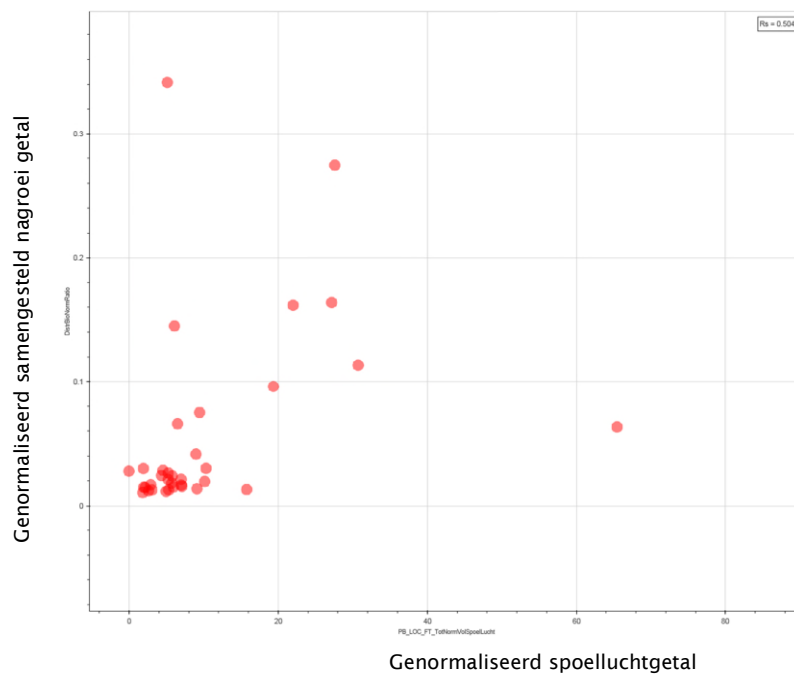
3.5.3 Spoelwaterpercentage en spoelluchtpercentage

Figuur 10 toont het positieve verband tussen het spoelwaterpercentage en het genormaliseerd KG22 van het gedistribueerde water (let op: het spoelwaterpercentage staat op de Y-as, het KG22 staat op de X-as). De Spearman coëfficiënt is 0.48.



FIGUUR 10. SPOELWATERPERCENTAGE (%) (Y-AS) VERSUS GENORMALISEERD KG22 VOOR GEDISTRIBUEERD WATER (X-AS).

Figuur 11 toont een vergelijkbaar positief verband tussen het spoelluchtpercentage en het genormaliseerd samengestelde nagroei getal van het gedistribueerde water. De Spearman coëfficiënt is 0,50. Deze observatie vormt een aanwijzing dat een lager spoelwater- of spoelluchtpercentage mogelijk een positief effect heeft op het verlagen van de nagroei-indicatoren.



FIGUUR 11. GENORMALISEERD SAMENGESTELD NAGROEI GETAL VOOR GEDISTRIBUEERD WATER VERSUS GENORMALISEERD SPOELLUCHTGETAL.

In beide gevallen zou eerder een negatief verband verwacht worden: naarmate het spoelwater- of spoellucht percentage toeneemt is de spoeling verondersteld beter, wordt gewenst materiaal effectief uit het filter verwijderd, en komt dat ten goede aan de waterkwaliteit. Een hypothese voor het positieve verband is dat de productielocaties die op basis van de ruwwaterkwaliteit een omvangrijk spoelwaterprogramma nodig hebben ondanks dat omvangrijke programma water produceren met een hogere nagroei. Een andere verklaring is mogelijk dat het spoelwaterprogramma “te” omvangrijk is, waardoor bijvoorbeeld bepaald materiaal van chemische of biologische samenstelling van het filtermedium wordt losgemaakt. Dat zou negatief kunnen werken in het geval dat het losgemaakte materiaal niet effectief uit het filter wordt verwijderd, of wanneer het materiaal wel effectief wordt verwijderd maar het losmaken en de verwijdering ongewenst is, of dat door het spoelen geen goede opbouw van bacteriële biomassa plaatsvindt over het filter.

3.5.4 Resterend bedrijfsvoeringsgerelateerde parameters

Alle relaties tussen de nagroei indicatoren in reinwater of het gedistribueerde water en bedrijfsvoeringsgerelateerde eigenschappen die een redelijk sterke correlatie coëfficiënt tonen én die een praktisch belang hebben, zijn hierboven benoemd. Er zijn geen sterke correlaties gevonden tussen nagroei indicatoren en de andere bedrijfsvoeringsgerelateerde eigenschappen. Een samenvattend overzicht van de bedrijfsvoeringsgerelateerde parameters is weergegeven in Tabel 6, en Bijlage I toont een completer overzicht. Uit die overzichten kan worden afgeleid met welke bedrijfsvoeringsgerelateerde eigenschappen in voorliggende studie geen relaties aangetoond zijn met de nagroei parameters. Uit dat volledige overzicht zijn hierna verschillende bedrijfsvoeringsgerelateerde geselecteerd waarvan op voorhand de verwachting bestond dat er mogelijk een correlatie bestaat met de nagroei indicatoren. In deze studie is geen verband gevonden tussen de volgende bedrijfsvoeringsgerelateerde parameters en nagroei indicatoren:

- Leeftijd filtermedium. Een ouder filtermedium zou mogelijk minder goed presteren vanwege slijtage, irreversibele ongewenste aangroei, of “dode” zones. Aan de andere kant is bekend dat filters een tijd moeten inlopen alvorens ze goed presteren.
- Aantal pompputschakelingen en aantal filterschakelingen per dag. Een groter aantal schakelingen zou kunnen leiden tot een grotere deeltjeslast in het reinwater en het gedistribueerde water, wat nadelig is voor de waterkwaliteit.
- Benuttingsgraad locatie (ratio gemiddeld en ontwerp capaciteit). Een hoge benuttingsgraad, of zelfs een waarde > 1, en een te lage benuttingsgraad zouden mogelijk nadelig kunnen zijn voor de waterkwaliteit, bijvoorbeeld wegens doorslag van deeltjes (hoge ratio) of voorkeursstroming in filters (te lage belasting).

Aan de variabiliteit van de bedrijfsvoering wordt verder aandacht besteed in paragraaf 3.6.4.

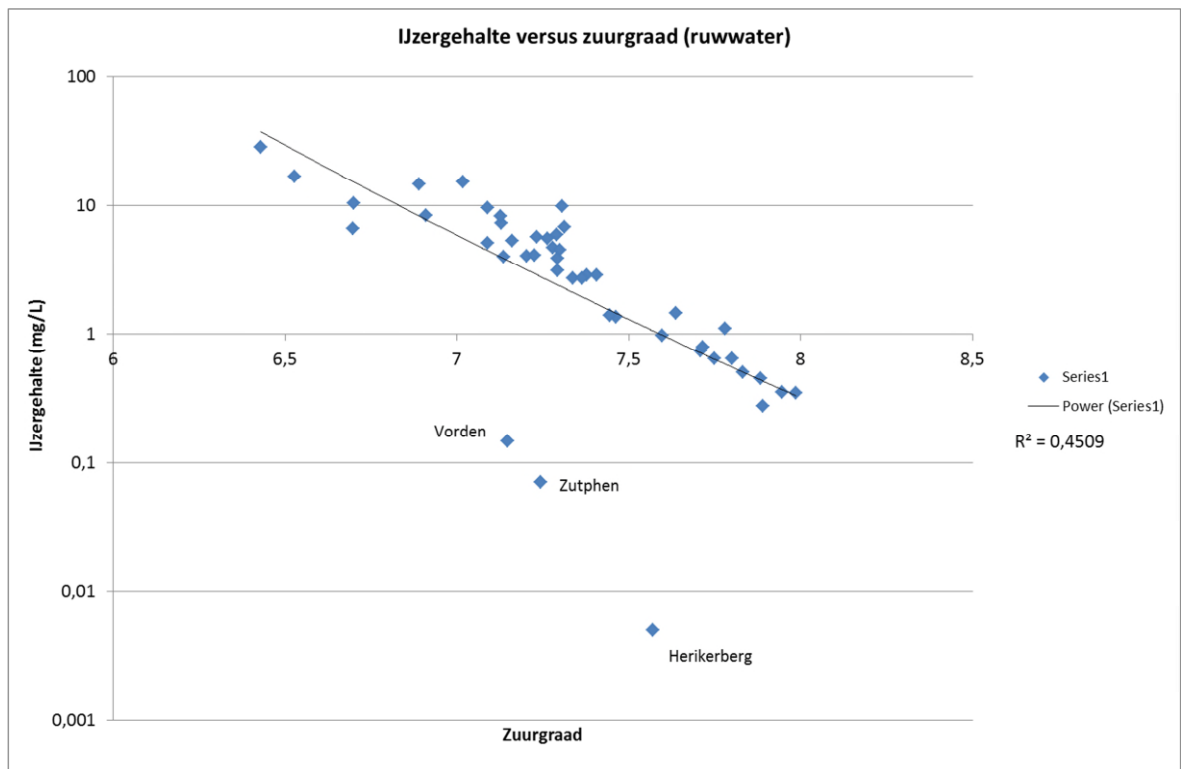
De afwezigheid van (één op één) relaties hoeft niet te betekenen dat parameters definitief geen invloed hebben op nagroei-indicatoren. De relatie kan door verstoringen door andere parameter(s) zijn weggevallen of is in combinatie met andere parameters (meervoudige regressie) wel van belang.

3.6 Hypothese-gedreven relaties nagroei indicatoren

In deze paragraaf is gepoogd om verschillende hypothesen kwalitatief te testen, in de meeste gevallen aan de hand van visualisatie middels 4D-plots. In deze 4D-figuren zijn naast de X-as en de Y-as parameter twee andere parameters toegevoegd door variatie van de kleur en de grootte van de indicator. Voor wat betreft de grootte van de indicator geldt dat de waarde van de parameters toeneemt naarmate de omvang van de indicator toeneemt. Voor wat betreft de kleur geldt dat rood overeenkomt met een hoge parameterwaarde en dat donkergroen overeenkomt met lage parameterwaarden (gehele kleurschakering van hoog naar laag is: rood – oranje – geel – lichtgroen – donkergroen).

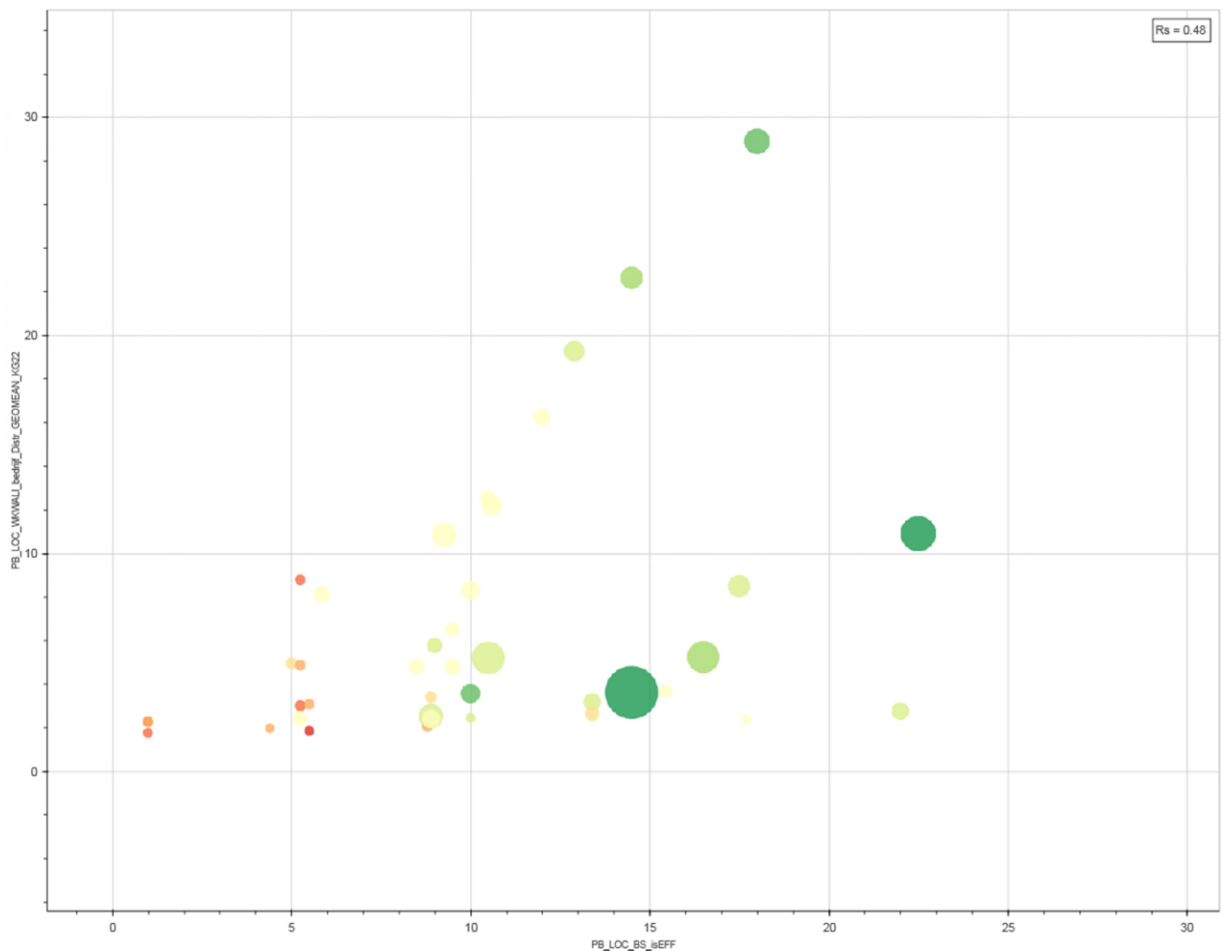
3.6.1 Hypothese biologische ontijzering

Voor de betrokken productielocaties is een verband gevonden tussen het ijzergehalte en de zuurgraad van het ruwwater, zie Figuur 12. De correlatie coëfficiënt (R^2) is 0,45, en exclusief de drie locaties met het laagste ijzergehalte is de betreffende coëfficiënt 0,85.



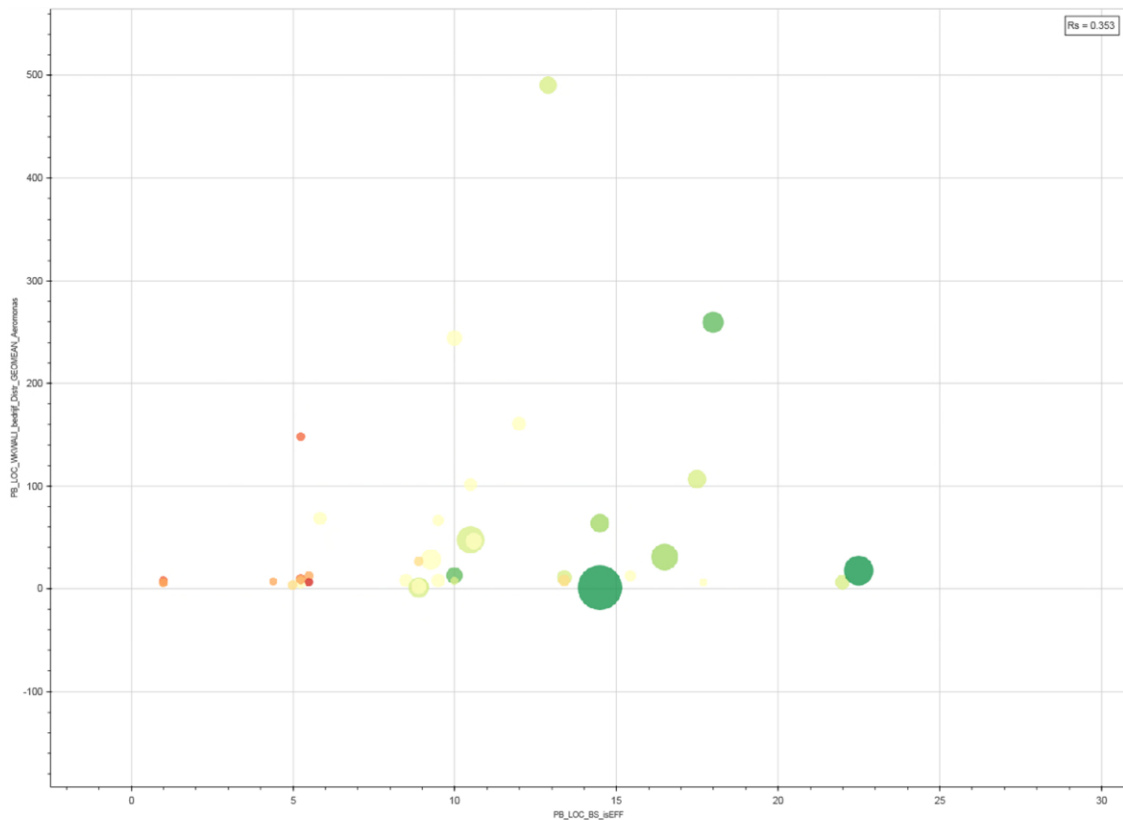
FIGUUR 12. VERBAND TUSSEN IJZERGEHALTE (MG/L) EN ZUURGRAAD (-) RUWWATER

Een hypothese is dat de ontijzering (deels) biologisch plaatsvindt bij een samenstelling van het ruwwater die gekenmerkt wordt door een hoog ijzergehalte en een lage pH, en dat de nagroei indicatoren hoger worden naarmate de ontijzering meer biologisch van aard is. Andere kenmerken van een zuivering waarbij de ontijzering biologisch kan verlopen zijn een hoge filtratiesnelheid en een beperkt zuurstofgehalte/beperkte beluchting. Om deze hypothese te testen is een visualisatie gemaakt waarin het ijzergehalte, de pH, het effectieve beluchtingsgetal en een nagroei indicator gecombineerd worden. Figuur 13 geeft deze visualisatie weer voor het KG22 van het gedistribueerde drinkwater.



FIGUUR 13. KG22 (KVE/ML) GEDISTRIBUEERD WATER VERSUS EFFECTIEF BELUCHTINGSGETAL VOOR VERSCHILLENDE RUWWATERSOORTEN: HET IJZERGEHALTE IS WEERGEGEVEN AAN DE HAND VAN DE GROOTTE VAN DE INDICATOR (GROTERE PUNT = HOGER IJZERGEHALTE; RANGE = 0,005 - 28,2 MG/L), EN DE PH IS WEERGEGEVEN AAN DE HAND VAN DE KLEUR VAN DE INDICATOR (ROOD = HOGE PH; GEEL = MIDDEN PH, GROEN = LAGE PH; RANGE = 6,4 - 8,0).

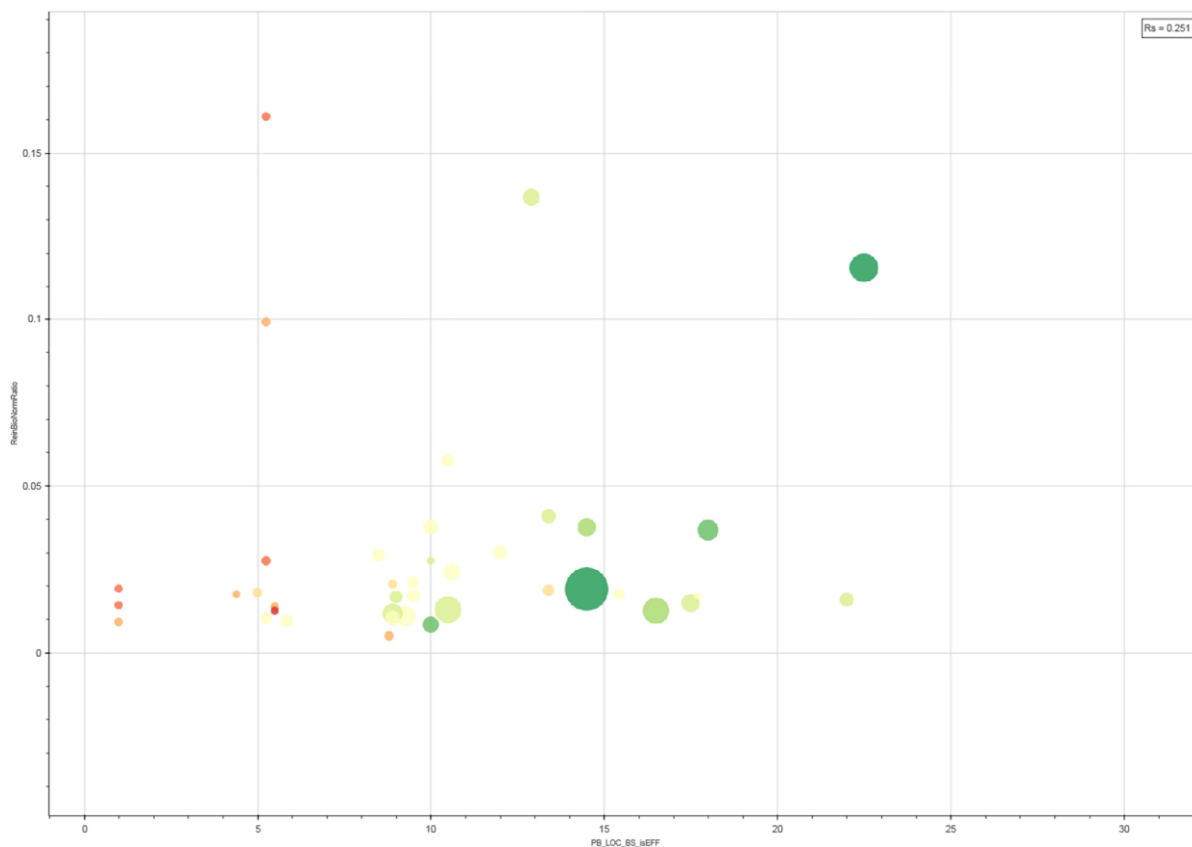
Uit Figuur 13 blijkt dat de locaties met een hoge pH en een lager ijzergehalte een relatief milde beluchting hebben (laag effectief beluchtingsgetal) omdat de pH niet sterk verhoogd dient te worden en weinig zuurstof nodig is voor de oxidatie van ijzer (zie kleine, rode en oranje punten links in de figuur). Bij deze locaties is de kans op / de geschiktheid voor biologische ontijzering klein. Deze locaties hebben lagere KG22 waarden in het gedistribueerde water. De locaties met een lagere pH en een hoger ijzergehalte (gele en groene indicatoren) hebben meer beluchting nodig om de pH te verhogen en om de componenten te oxideren, deze locaties hebben doorgaans een hoger effectief beluchtingsgetal. Bij deze locaties is de kans op / geschiktheid voor biologische ontijzering groter. Deze locaties hebben in sommige gevallen hogere KG22 waarden in het gedistribueerde water. Echter, de figuur toont ook locaties die mogelijk (gedeeltelijk) biologisch ontijzeren waarbij het KG22 van het gedistribueerde water lager is.



FIGUUR 14. *AEROMONAS* (KVE/100 ML) GEDISTRIBUEERD WATER VERSUS EFFECTIEF BELUCHTINGSGETAL VOOR VERSCHILLENDE RUWWATERSOORTEN: HET IJZERGEHALTE IS WEERGEGEVEN AAN DE HAND VAN DE GROOTTE VAN DE INDICATOR (GROTERE PUNT = HOGER IJZERGEHALTE; RANGE = 0,005 - 28,2 MG/L), EN DE PH IS WEERGEGEVEN AAN DE HAND VAN DE KLEUR VAN DE INDICATOR (ROOD = HOGE PH; GEEL = MIDDEN PH, GROEN = LAGE PH; RANGE = 6,4 - 8,0).

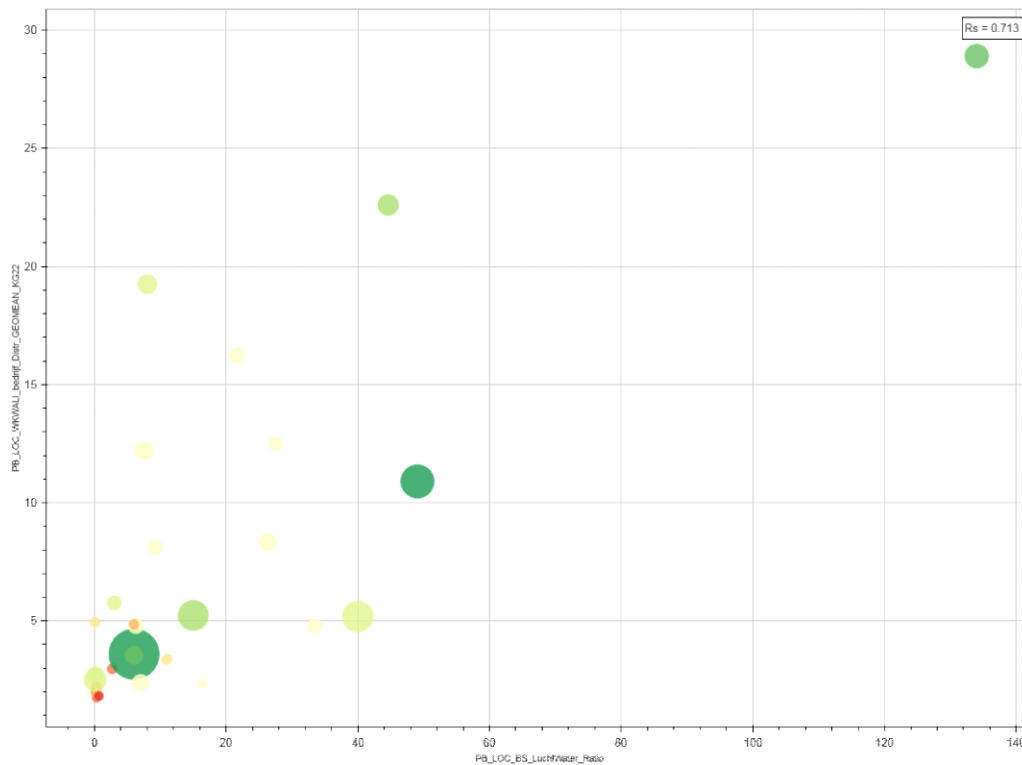
Een vergelijkbare observatie blijkt uit Figuur 14, waar het *Aeromonas* getal van het gedistribueerde water is geplot. Hieruit volgt een vergelijkbaar beeld voor de locaties met een hoge pH en een laag ijzergehalte, deze locaties hebben vaak een lager *Aeromonas* getal. De locaties met een hoger aantal *Aeromonas* zijn meestal locaties die mogelijk (gedeeltelijk) biologisch ontijzeren. Echter, de figuur toont ook locaties die mogelijk (gedeeltelijk) biologisch ontijzeren waarbij het *Aeromonas* aantal van het gedistribueerde water lager is.

Figuur 15 toont de waarde van de samengesteld genormaliseerde nagroei indicator voor het reinwater. De waarden zijn voor de meeste productielocaties relatief laag. Bij de locaties met een hogere waarde van de nagroei indicator zou de ontijzering (deels) biologisch kunnen verlopen. Dit is een aanwijzing dat de biologische ontijzering mogelijk leidt tot hogere nagroei indicatoren. Bij twee locaties met hogere waarden van de nagroei indicator is de kans op biologische ontijzering echter laag vanwege een hoge pH en een laag ijzergehalte van het ruwwater. Middels kwantificering van de biologische ontijzering middels qPCR of RNA analyse kan hier meer inzicht in verkregen worden.



FIGUUR 15. SAMENGESTELD GENORMALISEERD NAGROEI GETAL VOOR REINWATER VERSUS EFFECTIEF BELUCHTINGSGETAL VOOR VERSCHILLENDE RUWWATERSOORTEN: HET IJZERGEHALTE IS WEERGEGEVEN AAN DE HAND VAN DE GROOTTE VAN DE INDICATOR (GROTERE PUNT = HOGER IJZERGEHALTE; RANGE = 0,005 - 28,2 MG/L), EN DE PH IS WEERGEGEVEN AAN DE HAND VAN DE KLEUR VAN DE INDICATOR (ROOD = HOGE PH; GEEL = MIDDEN PH, GROEN = LAGE PH; RANGE = 6,4 - 8,0).

Uit de plots met de lucht-water verhouding in plaats van het effectief beluchtingsgetal volgt een vergelijkbaar beeld, zie Figuur 16. Uit deze figuur kan tussen het vergelijk van de locaties met een hoog ijzergehalte en een lage pH (groene, grotere punten) nog worden afgeleid dat de locaties die op basis van de ruwwaterkwaliteit mogelijk geschikt zijn voor biologische ontijzering én ook een lage lucht-water ratio hebben beter presteren dan de locaties met vergelijkbare ruwwaterkwaliteit (Fe en pH) en een hoge lucht-water ratio. Reeds eerder (paragraaf 3.5.1) is het positieve verband tussen de nagroei indicatoren en de lucht-water verhouding beschreven.

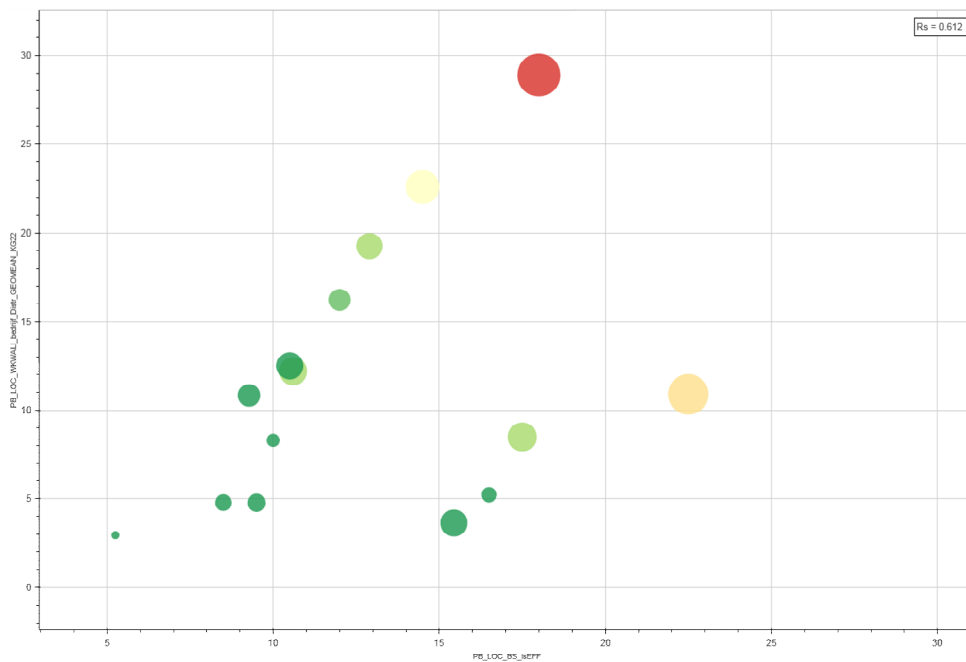


FIGUUR 16. GENORMALISEERD KG22 VOOR GEDISTRIBUEERD WATER VERSUS LUCHT-WATER RATIO VOOR VERSCHILLENDE RUWWATERSOORTEN: HET IJZERGEHALTE IS WEERGEGEVEN AAN DE HAND VAN DE GROOTTE VAN DE INDICATOR (GROTERE PUNT = HOGER IJZERGEHALTE; RANGE = 0,005 - 28,2 MG/L), EN DE PH IS WEERGEGEVEN AAN DE HAND VAN DE KLEUR VAN DE INDICATOR (ROOD = HOGE PH; GEEL = MIDDEN PH, GROEN = LAGE PH; RANGE = 6,4 - 8,0).

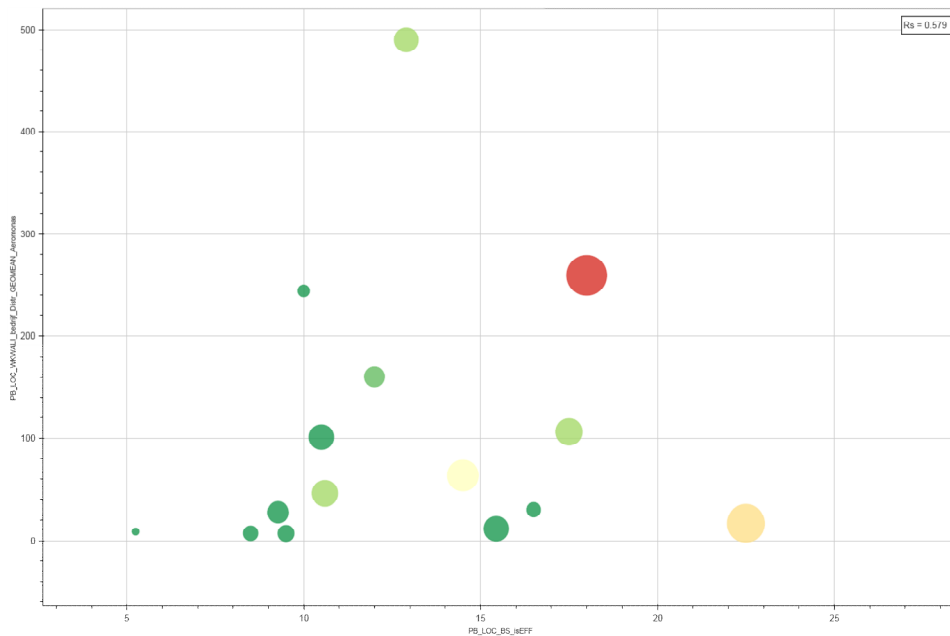
3.6.2 Hypothese ruwwater samenstelling, beluchting en nagroei indicator

In voorliggend en voorgaand onderzoek zijn correlaties gevonden tussen het TOC en methaan gehalte van ruwwater en nagroei indicatoren in het gedistribueerde water. Deze bevindingen worden hier aan de hand van verschillende figuren gevisualiseerd door een nagroei indicator, het effectief beluchtingsgetal, het methaan en het TOC gehalte van het ruwwater te combineren in één visualisatie.

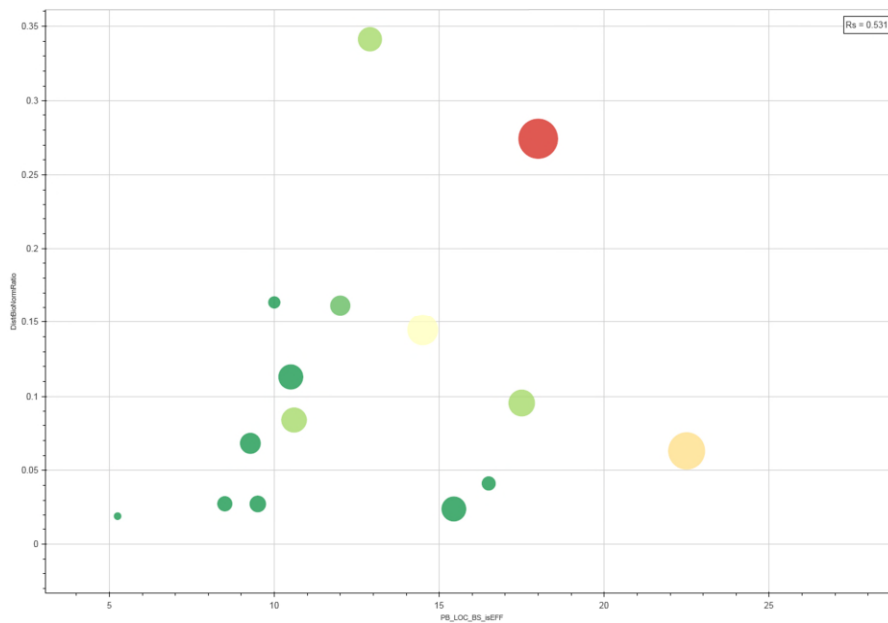
De vier plots (Figuur 17 - Figuur 20) tonen het te verwachten beeld dat de locaties met een lager TOC of methaan gehalte (de kleinere, groenere punten) een minder intensieve beluchting nodig hebben en dat deze locaties vaker lagere waarden van de nagroei indicatoren hebben dan de locaties met een hoger TOC of methaan gehalte (de lichtgroene, gele en rode grotere punten).



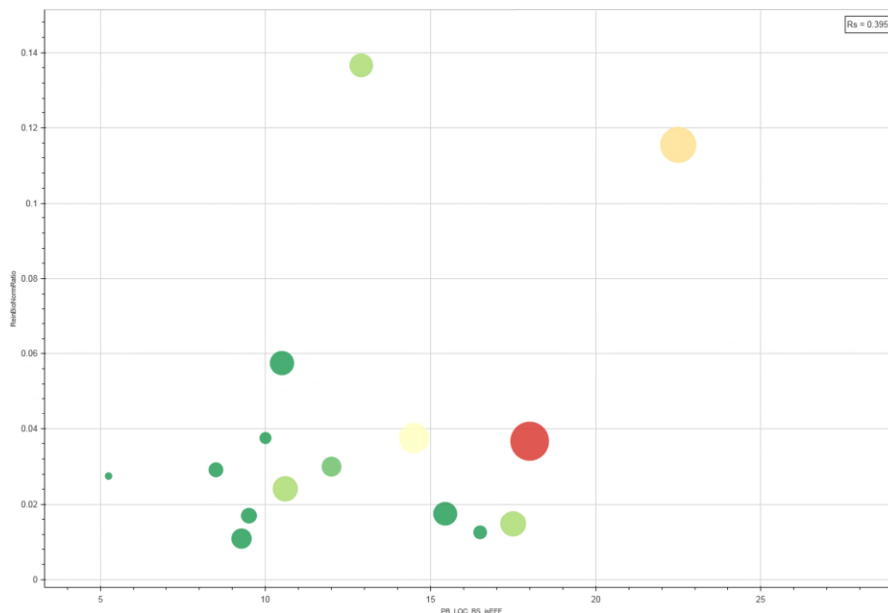
FIGUUR 17. KOLONEGETAL GEDISTRIBUEERD WATER VERSUS EFFECTIEF BELUCHTINGSGETAL VOOR VERSCHILLENDE RUWWATERSOORTEN: HET TOC GEHALTE IS WEERGEGEVEN AAN DE HAND VAN DE GROOTTE VAN DE INDICATOR (GROTERE PUNT = HOGER TOC GEHALTE; RANGE = 0,25 – 7,9 MG/L), EN HET METHAAN GEHALTE IS WEERGEGEVEN AAN DE HAND VAN DE KLEUR VAN DE INDICATOR (ROOD = HOOG METHAAN; GEEL = MIDDEN METHAAN, GROEN = LAAG METHAAN; RANGE = 0,005 – 18,9 MG/L).



FIGUUR 18. AEROMONAS GEDISTRIBUEERD WATER VERSUS EFFECTIEF BELUCHTINGSGETAL VOOR VERSCHILLENDE RUWWATERSOORTEN: HET TOC GEHALTE IS WEERGEGEVEN AAN DE HAND VAN DE GROOTTE VAN DE INDICATOR (GROTERE PUNT = HOGER TOC GEHALTE; RANGE = 0,25 – 7,9 MG/L), EN HET METHAAN GEHALTE IS WEERGEGEVEN AAN DE HAND VAN DE KLEUR VAN DE INDICATOR (ROOD = HOOG METHAAN; GEEL = MIDDEN METHAAN, GROEN = LAAG METHAAN; RANGE = 0,005 – 18,9 MG/L).



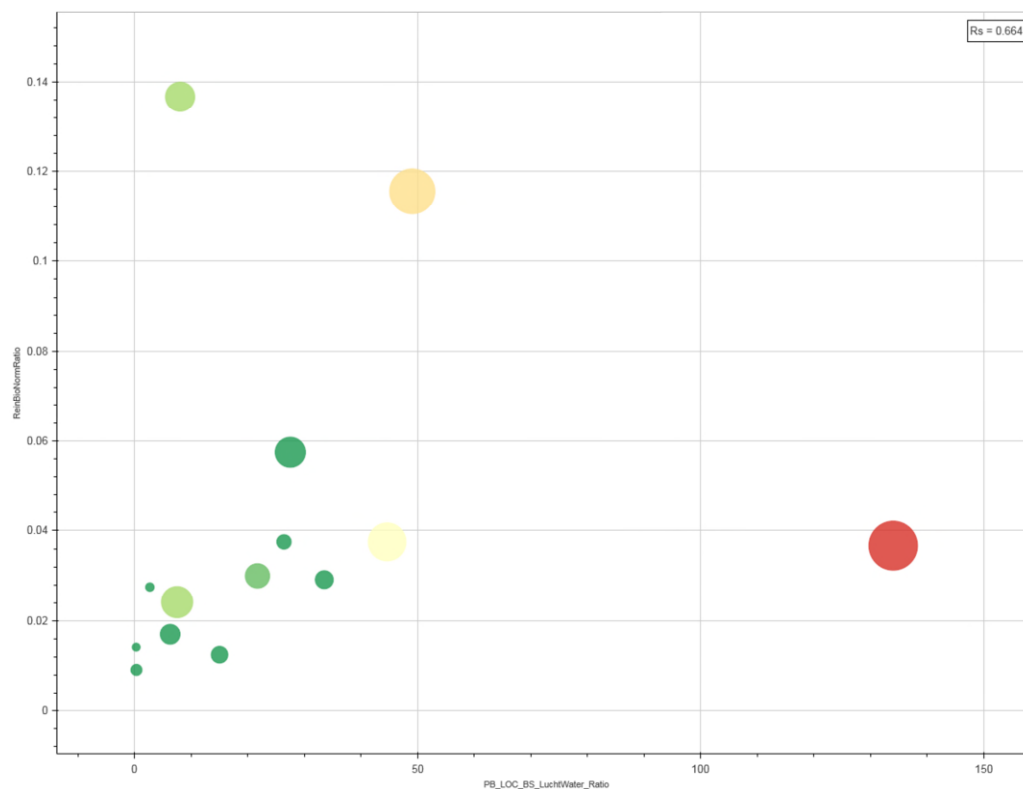
FIGUUR 19. SAMENGESTELD, GENORMALISEERD NAGROEI GETAL GEDISTRIBUEERD WATER VERSUS EFFECTIEF BELUCHTINGSGETAL VOOR VERSCHILLENDE RUWWATERSOORTEN: HET TOC GEHALTE WEERGEGEVEN AAN DE HAND VAN GROOTTE VAN INDICATOR (GROTERE PUNT = HOGER TOC GEHALTE; RANGE = 0,25 - 7,9 MG/L), EN METHAAN GEHALTE WEERGEGEVEN AAN DE HAND VAN KLEUR VAN INDICATOR (ROOD = HOOG METHAAN; GEEL = MIDDEN METHAAN, GROEN = LAAG METHAAN; RANGE = 0,005 - 18,9 MG/L).



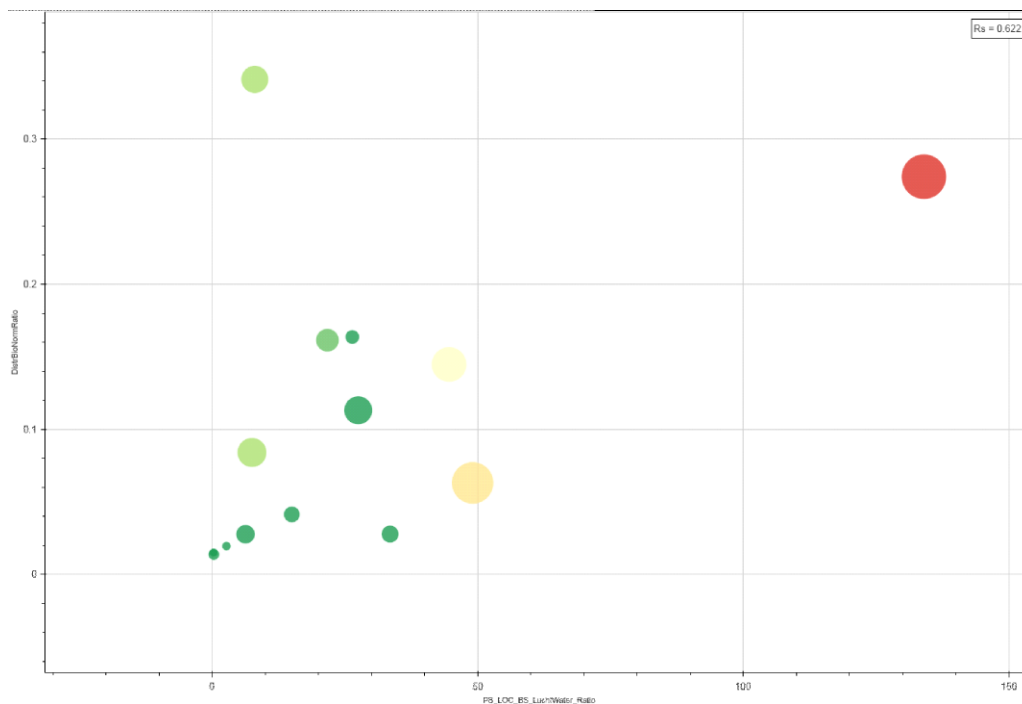
FIGUUR 20. SAMENGESTELD, GENORMALISEERD NAGROEI GETAL REINWATER VERSUS EFFECTIEF BELUCHTINGSGETAL VOOR VERSCHILLENDE RUWWATERSOORTEN: TOC GEHALTE WEERGEGEVEN AAN DE HAND VAN GROOTTE VAN INDICATOR (GROTERE PUNT = HOGER TOC GEHALTE; RANGE = 0,25 - 7,9 MG/L), EN METHAAN GEHALTE WEERGEGEVEN AAN DE HAND VAN DE KLEUR VAN DE INDICATOR (ROOD = HOOG METHAAN; GEEL = MIDDEN METHAAN, GROEN = LAAG METHAAN; RANGE = 0,005 - 18,9 MG/L).

Vergelijkbare visualisaties zijn opgesteld met de nagroei indicatoren, de lucht-water verhouding (in plaats van het effectief beluchtingsgetal), het methaan en het TOC gehalte van het ruwwater. Figuur 21 en Figuur 22 tonen deze combinatie voor het samengesteld genormaliseerd nagroei getal voor het rein- en voor het gedistribueerde water. Uit deze figuren volgt het beeld dat de locaties met een lager methaan gehalte (groenere punten) een minder intensieve beluchting hebben en vaker lagere aantallen van de nagroei indicatoren laten zien.

Dit beeld onderschrijft de eerder beschreven relaties. Een combinatie van het beluchtingstype, het effectief beluchtingsgetal, de lucht-water ratio, de methaan en TOC concentratie in ruwwater lijkt een relatie te hebben met de mate van nagroei die in het gedistribueerde water optreedt. Aanvullend onderzoek is nodig om te bepalen welke van deze parameters het meest belangrijk zijn in die relatie.



FIGUUR 21. SAMENGESTELD, GENORMALISEERD NAGROEI GETAL REINWATER VERSUS LUCHT-WATER RATIO VOOR VERSCHILLENDE RUWWATERSOORTEN: HET TOC GEHALTE IS WEERGEGEVEN AAN DE HAND VAN DE GROOTTE VAN DE INDICATOR (GROTERE PUNT = HOGER TOC GEHALTE; RANGE = 0,25 - 7,9 MG/L), EN HET METHAAN GEHALTE IS WEERGEGEVEN AAN DE HAND VAN DE KLEUR VAN DE INDICATOR (ROOD = HOOG METHAAN; GEEL = MIDDEN METHAAN, GROEN = LAAG METHAAN; RANGE = 0,005 - 18,9 MG/L).

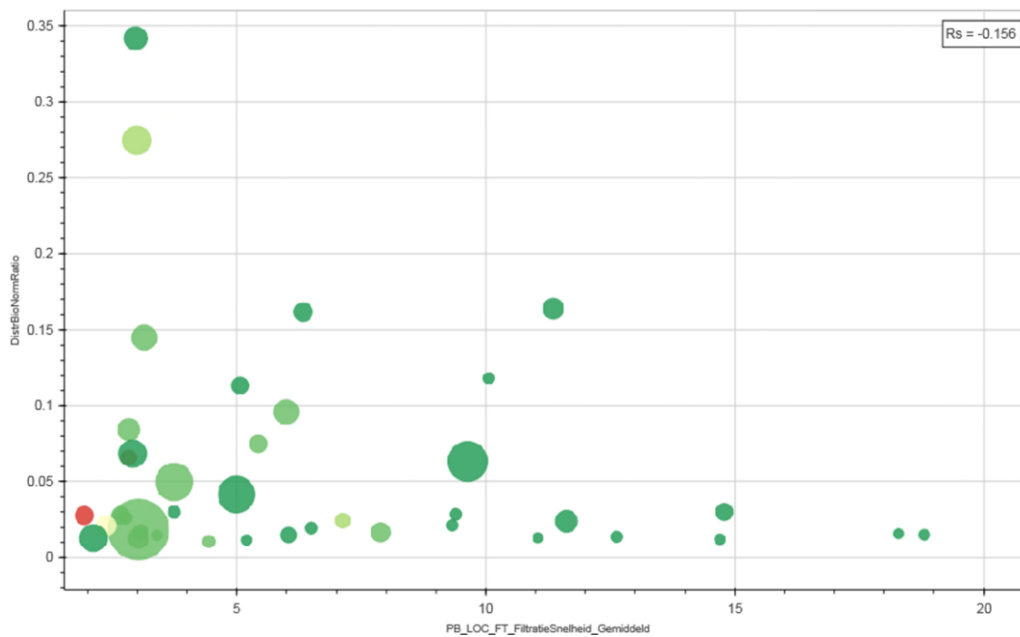


FIGUUR 22. SAMENGESTELD, GENORMALISEERD NAGROEI GETAL GEDISTRIBUEERD WATER VERSUS LUCHT-WATER RATIO VOOR VERSCHILLENDE RUWWATERSOORTEN: HET TOC GEHALTE IS WEERGEGEVEN AAN DE HAND VAN DE GROOTTE VAN DE INDICATOR (GROTERE PUNT = HOGER TOC GEHALTE; RANGE = 0,25 - 7,9 MG/L), EN HET METHAAN GEHALTE IS WEERGEGEVEN AAN DE HAND VAN DE KLEUR VAN DE INDICATOR (ROOD = HOOG METHAAN; GEEL = MIDDEN METHAAN, GROEN = LAAG METHAAN; RANGE = 0,005 - 18,9 MG/L).

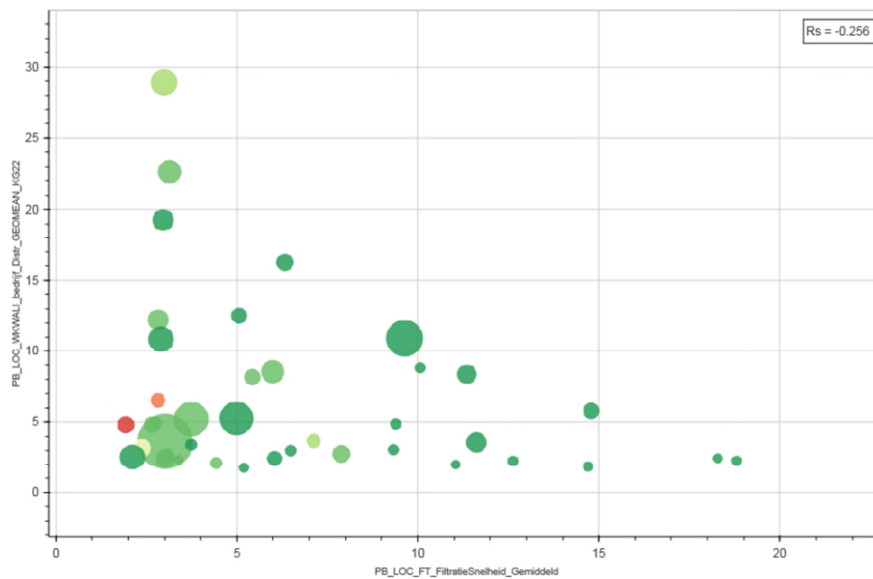
3.6.3 Hypothese filterbelasting

Een hypothese is dat filters die hoger belast worden door hoge concentraties van grondwaterparameters in het ruwwater (in dit voorbeeld ijzer en ammonium) op lagere snelheid bedreven worden en dat er een optimum snelheid bestaat: een te lage snelheid leidt tot voorkeursstroming en een te hoge snelheid leidt tot doorslag of onvolledige verwijdering.

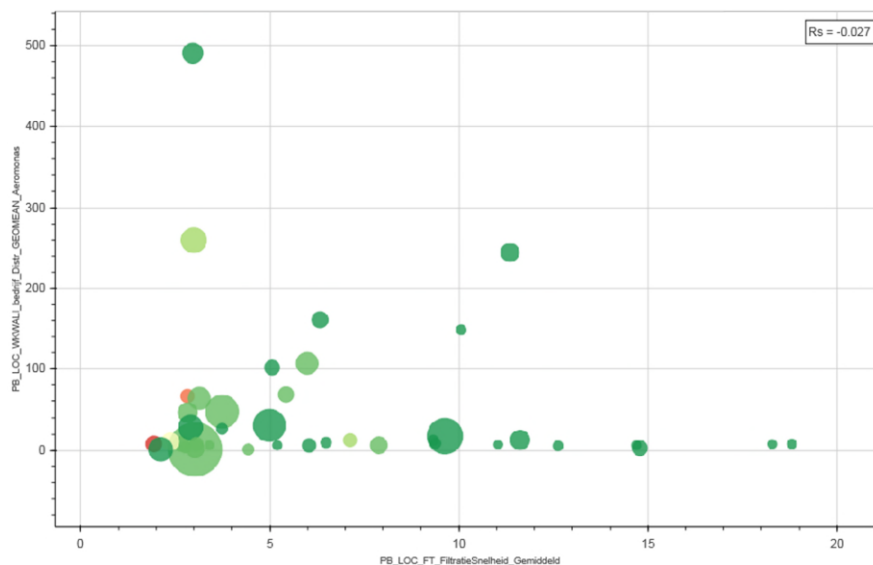
Figuur 23 laat zien dat locaties met een hoog ammonium gehalte vaak op een lagere filtratiesnelheid bedreven worden. De locaties met een lager ijzer en een lager ammonium gehalte (kleine, groene punten) worden op hogere filtratiesnelheden bedreven. De hoogste waarden van de nagroei indicatoren komen voor bij locaties met een lage filtratiesnelheid, zonder dat bij deze locaties sprake is van zeer hoge ijzer of ammonium waarden. Figuur 24 en Figuur 25 tonen met de achterliggende waarden van Figuur 23 een verdeelder beeld. Uit deze figuren kan niet worden afgeleid op welke wijze de filtratiesnelheid van invloed is op de nagroei-indicatoren, echter het beeld toont dat geen sprake is van een optimale filtratiesnelheid.



FIGUUR 23. SAMENGESTELD, GENORMALISEERD NAGROEI GETAL GEDISTRIBUEERD WATER VERSUS FILTRATIESNELHEID (M/H) VOOR VERSCHILLENDE RUWWATERSOORTEN: HET IJZER GEHALTE IS WEERGEGEVEN AAN DE HAND VAN DE GROOTTE VAN DE INDICATOR (GROTERE PUNT = HOGER IJZER GEHALTE; RANGE = 0,005 - 28 MG/L), EN HET AMMONIUM GEHALTE IS WEERGEGEVEN AAN DE HAND VAN DE KLEUR VAN DE INDICATOR (ROOD = HOOG AMMONIUM; GEEL = MIDDEN AMMONIUM, GROEN = LAAG AMMONIUM; RANGE = 0,015 - 5,5 MG/L).



FIGUUR 24. GENORMALISEERD KG22 GEDISTRIBUEERD WATER VS. FILTRATIESNELHEID (M/H) VOOR VERSCHILLENDE RUWWATERSOORTEN: IJZER GEHALTE IS WEERGEGEVEN AAN DE HAND VAN GROOTTE VAN INDICATOR (GROTERE PUNT = HOGER IJZER GEHALTE; RANGE = 0,005 - 28 MG/L), EN HET AMMONIUM GEHALTE IS WEERGEGEVEN AAN DE HAND VAN DE KLEUR VAN DE INDICATOR (ROOD = HOOG AMMONIUM; GEEL = MIDDEN AMMONIUM, GROEN = LAAG AMMONIUM; RANGE = 0,015 - 5,5 MG/L).



FIGUUR 25. GENORMALISEERD *AEROMONAS* GEDISTRIBUEERD WATER VS. FILTRATIESNELHEID (M/H) VOOR VERSCHILLENDE RUWWATERSOORTEN: IJZER GEHALTE IS WEERGEGEVEN AAN DE HAND VAN GROOTTE VAN DE INDICATOR (GROTERE PUNT = HOGER IJZER GEHALTE; RANGE = 0,005 - 28 MG/L), EN HET AMMONIUM GEHALTE IS WEERGEGEVEN AAN DE HAND VAN DE KLEUR VAN DE INDICATOR (ROOD = HOOG AMMONIUM; GEEL = MIDDEN AMMONIUM, GROEN = LAAG AMMONIUM; RANGE = 0,015 - 5,5 MG/L).

3.6.4 Hypothese variabiliteit bedrijfsvoering

Een groot aantal ontwerp- en bedrijfsvoeringsgerelateerde eigenschappen is gerelateerd aan de variabiliteit van de productie. De verwachting was dat een constante bedrijfsvoering gunstiger is voor de waterkwaliteit. Eigenschappen die hierop betrekking hebben zijn:

- Wordt het snelfilter na een spoeling, na een schakeling of na inlopen meteen of geleidelijk in bedrijf genomen?
- Worden filters dagelijks of wekelijks geschakeld, en zo ja hoe vaak? Is er sprake van carousel bedrijf?
- Worden pompputten regelmatig geschakeld, en zo ja, hoe vaak?
- Vinden debietsvariaties plaats?

Uit de enkelvoudige relaties tussen deze eigenschappen en de waterkwaliteit blijkt dat de Spearman coëfficiënten laag zijn, met uitzondering van de correlatie coëfficiënt met ijzer in het gedistribueerde water en het aantal filter schakelingen per dag (0,66). In het verleden is bij productielocatie Nuland ook waargenomen dat een continue bedrijfsvoering leidde tot een lagere ijzerafzettingssnelheid dan een discontinue bedrijfsvoering (van der Wielen, 2013).

Een aanvullende analyse op deze variabiliteit is uitgevoerd door de relatie tussen het aantal parallelle filterunits en verschillende waterkwaliteitsparameters te controleren. Daarbij is het idee dat het effect van een verstoring, zoals een schakeling of een spoeling, op de waterkwaliteit kleiner is bij een locatie met een groter aantal parallelle units. Uit deze analyse blijkt dat de locaties met een groter aantal parallelle units (6 van de 43 locaties met 10 of meer parallelle units) allen een laag KG22 en een laag *Aeromonas* aantal in het gedistribueerde water hebben ten opzichte van de overige locaties. Dit geldt tevens voor het ijzergehalte en de troebelingsgraad in het reinwater. Bij deze zes locaties is tevens sprake van een lage filtratiesnelheid. Mogelijk wordt het gunstige effect op de waterkwaliteit beïnvloed wordt door het aantal filterunits, al is in 3.6.3 getoond dat hoge nagroei indicator waarden ook voorkomen bij lage filtratiesnelheden.

3.7 Test invloed binaire variabelen op waterkwaliteit

In de weergave van het KG22 in het gedistribueerde water versus het effectief beluchtingsgetal, maar ook versus andere parameters zoals het ijzer- en mangaangehalte in ruwwater, lijken zich twee groepen of clusters te vormen. Naar aanleiding van deze observatie is voor het KG22 en *Aeromonas* in het gedistribueerde water en het ijzergehalte en de troebelingsgraad van het reinwater onderzocht of de waarde van de waterkwaliteit parameter verklaard wordt door een continue parameter en een binaire parameter (TRUE/FALSE). Op deze wijze ontstaat als het ware een 3D-visualisatie. De werkwijze is beschreven in paragraaf 2.6.4. Voor het KG22 levert deze werkwijze enkele interessante aanwijzingen op. Deze worden in Figuur 26 getoond. Voor de overige onderzochte parameters (*Aeromonas* in het gedistribueerde water, het ijzergehalte en de troebelingsgraad van het reinwater) zijn geen leads gevonden.

Uit Figuur 26 blijkt dat hogere waarden van het KG22 in het gedistribueerde water vaker voorkomt bij productielocaties waarbij sprake is van:

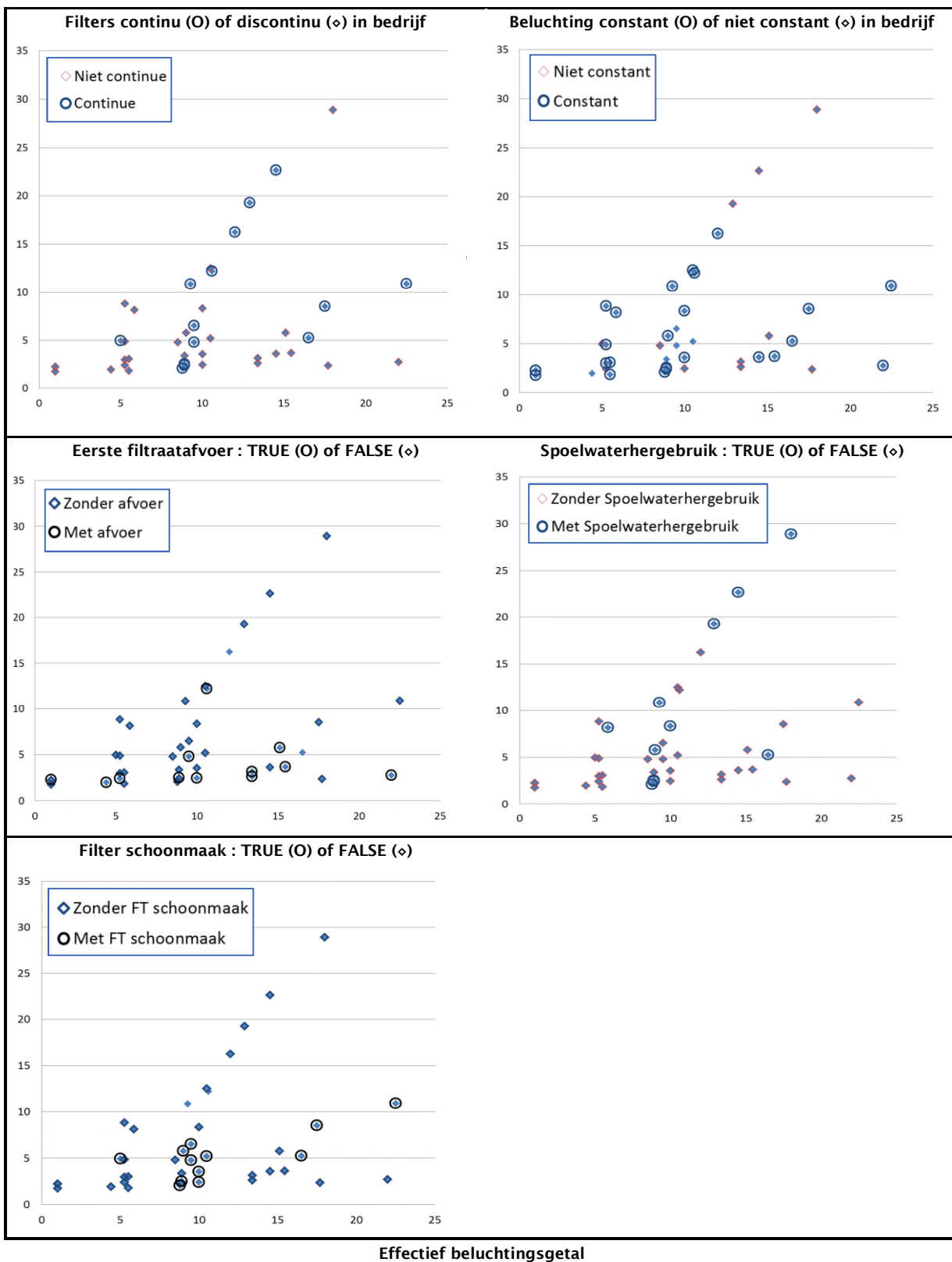
- Geen schoonmaak van filters;
- Geen eerste filtraat afvoer;
- Spoelwaterhergebruik;
- Een discontinue bedrijfsvoering van de beluchting, maar een continue bedrijfsvoering van de filters.

Dit zijn aanwijzingen dat bij productielocaties waarbij de filters niet worden schoongemaakt, geen sprake is van eerste filtraat afvoer, sprake is van spoelwaterhergebruik, waarvan de filters continue bedreven worden mogelijk sprake is van hogere nagroei in het gedistribueerde water. Deze bevindingen vormen mogelijk aanwijzingen om voor bestaande

locaties met hogere waarden van het KG22 te onderzoeken of optimalisatie op deze eigenschappen mogelijk is.

Om de mogelijke invloed van zulke ontwerp en bedrijfsvoeringsgerelateerde eigenschappen op de waterkwaliteit nader te onderzoeken kunnen verschillende statistische testen uitgevoerd worden, zoals de de Student T-test (toegepast wanneer de data normaal verdeeld is en een gelijke variantie heeft), de Welch test (toegepast wanneer de data normaal verdeeld is maar een ongelijke variantie heeft) of de Mann-Witney test (toegepast wanneer de data niet-normaal verdeeld populatie is). Deze testen zijn in voorliggende studie niet uitgevoerd.

KG22 (kve/mL) gedistribueerde water

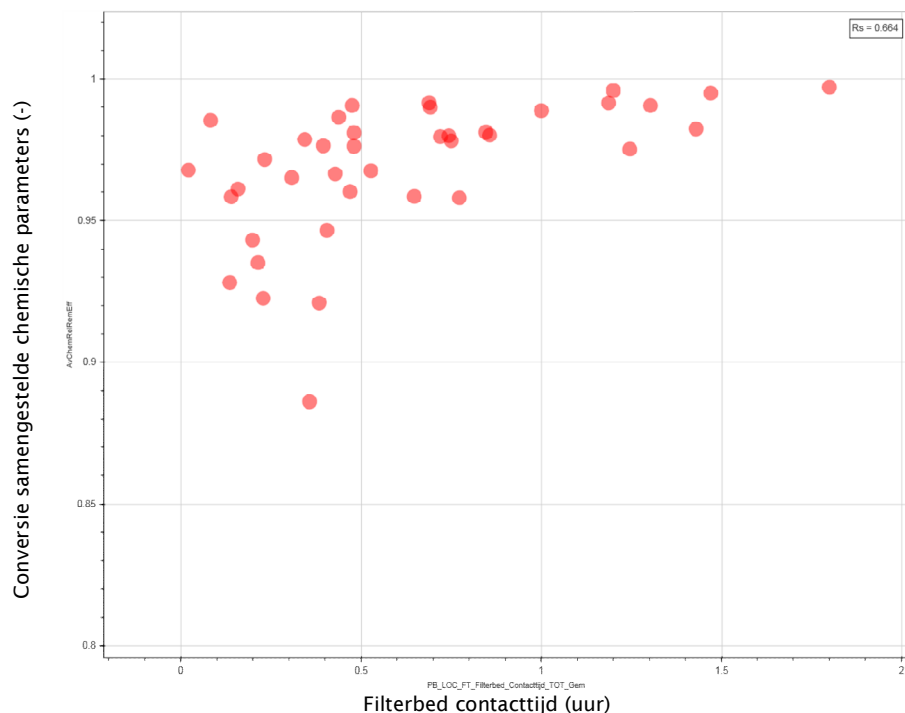


FIGUUR 26. KG22 GEDISTRIBUEERDE WATER VS EFFECTIEF BELUCHTINGSGETAL BIJ VERSCHILLENDE ONTWERP- EN BEDRIJFSVOERINGSEIGENSCHAPPEN.

3.8 Relaties chemische waterkwaliteit en ontwerp of bedrijfsvoering

In deze paragraaf worden de relaties tussen chemische waterkwaliteit en ontwerp- of bedrijfsvoeringsparameters besproken waarvan de Spearman coëfficiënten relatief hoog zijn en waarbij sprake is van praktische relevantie. Het uitgangspunt is hierbij is dat de aanwezigheid van bepaalde chemische (opgeloste of deeltjesgebonden) componenten een negatieve invloed kunnen hebben op de biologische stabiliteit van het reinwater en kunnen bijdragen aan nagroei in het leidingnet omdat deze stoffen kunnen dienen als voedingsstoffen voor micro-organismen en de deeltjes oppervlak bieden voor de groei van micro-organismen.

Figuur 27 toont het verwijderingsrendement (conversie) van de samengestelde chemische parameters ijzer, mangaan en ammonium versus de filterbed contacttijd (uur). De betreffende Spearman coëfficiënt is 0,66. Uit deze relatie blijkt dat de conversie van de chemische parameters toeneemt naarmate de filterbed contacttijd langer is. Vergelijkbare verbanden zijn ook gevonden voor de (afzonderlijke) absolute verwijdering en de conversie van TOC, ijzer, ammonium, en mangaan (met respectievelijke Spearman coëfficiënten voor de absolute verwijdering: 0.70, 0.61, 0.61, en 0.59). Een deel van de relaties met de afzonderlijke parameters is getoond in Bijlage II.

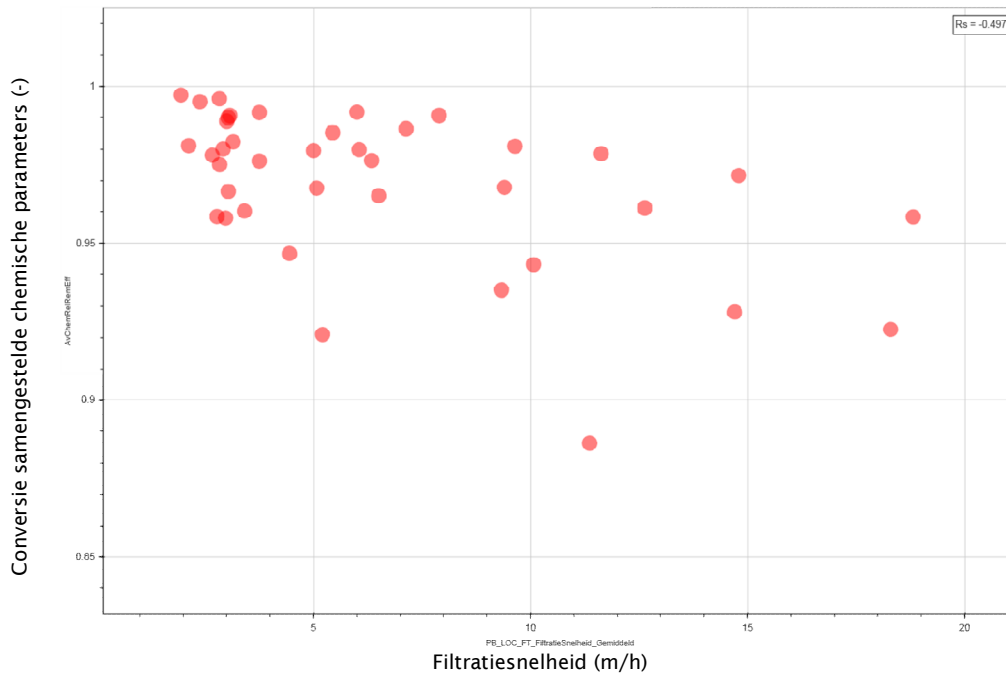


FIGUUR 27. CONVERSIE SAMENGESTELDE CHEMISCHE PARAMETERS (IJZER + MANGAAN + AMMONIUM) VERSUS FILTERBED CONTACTTIJD.

Figuur 28 toont het verwijderingsrendement (conversie) van de samengestelde chemische parameters ijzer, mangaan en ammonium versus de filtratiesnelheid (m/h). De betreffende Spearman coëfficiënt is - 0,50. De filtratiesnelheid is gerelateerd aan de filterbed contacttijd, maar er is geen sprake van een 1-op-1 relatie. Uit dit verband blijkt dat de conversie van de chemische parameters afneemt naarmate de filtratiesnelheid hoger is.

Vergelijkbare verbanden zijn gevonden voor de (afzonderlijke) absolute verwijdering van TOC en ammonium (met respectievelijke Spearman coëfficiënten: - 0.54 en - 0.65), de

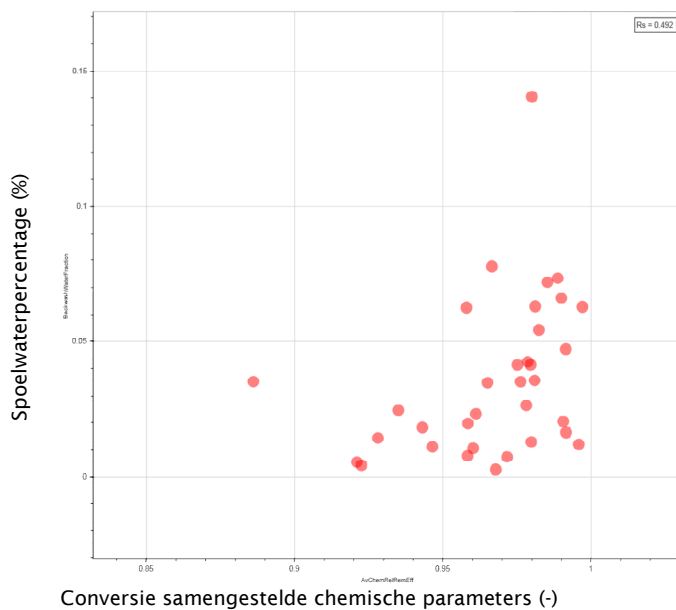
verbanden met de conversie van ijzer en mangaan zijn zwakker. Een deel van de relaties met deze afzonderlijke parameters is getoond in Bijlage II.



FIGUUR 28. CONVERSIE SAMENGESTELDE CHEMISCHE PARAMETERS (IJZER + MANGAAN + AMMONIUM) VERSUS FILTERBED CONTACTTIJD.

Figuur 29 toon het verwijderingsrendement (conversie) van de samengestelde chemische parameters (ijzer, mangaan en ammonium; op X-as) versus het spoelwaterpercentage (%) op Y-as). De betreffende Spearman coëfficiënt is 0,50. Uit deze relatie blijkt dat de conversie van de chemische parameters toeneemt naarmate het spoelwaterpercentage toeneemt. Een vergelijkbaar verband is gevonden met het spoelluchtpercentage (Spearman coëfficiënt 0,43).

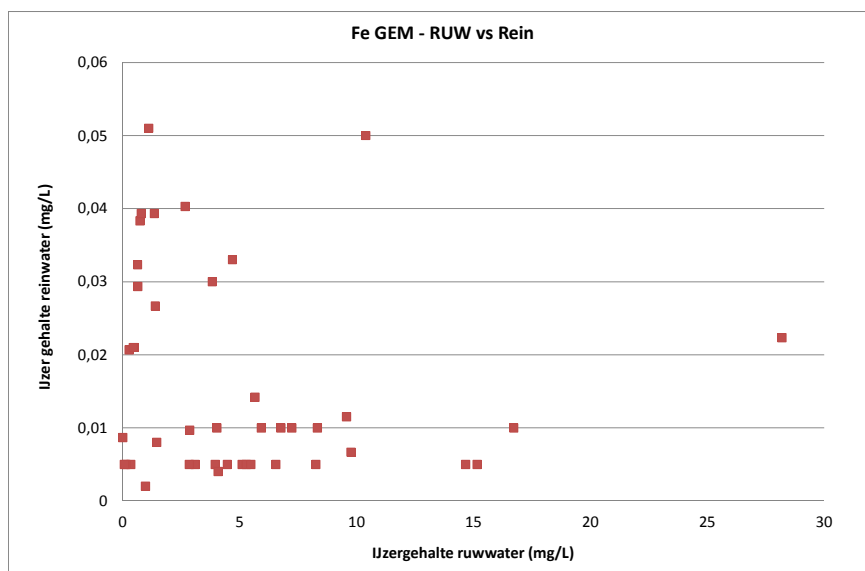
Vergelijkbare verbanden zijn ook gevonden voor de afzonderlijke absolute verwijdering van ijzer, TOC en mangaan (respectievelijke Spearman coëfficiënten 0.76, 0.6 en 0.56). De absolute verwijdering is gecorreleerd aan de ruwwaterkwaliteit. Een deel van de relaties met de afzonderlijke parameters is getoond in Bijlage II.



FIGUUR 29. CONVERSIE SAMENGESTELDE CHEMISCHE PARAMETERS (IJZER + MANGAAN + AMMONIUM) VERSUS SPOELWATERPERCENTAGE. LET OP: CONVERSIE OP X-AS EN SPOELWATERPERCENTAGE OP Y-AS.

Verschillende overige observaties:

- Er is een negatief verband gevonden tussen de verwijdering van ijzer en de benuttingsgraad van de locatie (Spearman coëfficiënt - 0,5). Dit zou er op kunnen duiden dat locaties die een grotere vracht aan ijzer dienen te verwijderen vaker onder de ontwerpcapaciteit bedreven worden.
- De hogere ijzergehalten in reinwater (tot 0,05 mg/L) komen voor bij de productielocaties die een laag ijzergehalte in het ruwwater hebben (Figuur 30). De richting van de relatie tussen ijzer in het reinwater en het gedistribueerde water is stijgend.



FIGUUR 30. IJZER REINWATER VERSUS IJZER RUWWATER

In Bijlage II zijn enkele additionele verbanden beschreven.

3.9 Resultaten clustering

Clustering is een techniek waarbij instanties in een dataset worden gegroepeerd in clusters, op basis van een afstandscriterium. Hoe groter de afstand tussen van een punt tot een andere groep (meet)punten, hoe groter de kans dat dit specifieke punt tot een apart cluster wordt gerekend. Deze techniek biedt de mogelijkheid om zuiveringen met overeenkomstige eigenschappen in een apart cluster in te delen, terwijl andere zuiveringen die andere effecten veroorzaken, zullen worden gegroepeerd in een ander cluster. Clusteranalyse kan een mogelijk startpunt zijn om zuiveringen in een specifiek cluster verder te analyseren. De clusteranalyses die in dit onderzoek zijn uitgevoerd hebben echter niet geleid tot een bruikbare indeling van productielocaties voor verdere analyse en zijn daarom niet beschreven in dit rapport.

4 Gevonden relaties en mogelijke hypothesen

In dit hoofdstuk zijn de gevonden relaties nogmaals benoemd en deels voorzien van een hypothese. Voor alle gevonden relaties geldt dat de aanwezigheid van een causaal verband niet onderzocht is. De genoemde relaties en hypothesen verklaren niet 100% van de variantie, echter zij geven mogelijke richtingen aan voor de optimalisatie van productielocaties en geven aanleiding tot gericht vervolgonderzoek. Sommige gevonden relaties voldoen aan het verwachtingspatroon, andere relaties geven een onverwacht beeld. Voor bepaalde gevonden verbanden zijn mogelijk achterliggende oorzaken benoemd, echter deze zijn in voorliggend project niet nader onderzocht

4.1 Relaties met nagroei-indicatoren

In deze paragraaf zijn de gevonden verbanden tussen de waarden van nagroei-indicatoren en de ontwerp- en bedrijfsvoeringsgerelateerde eigenschappen en overige waterkwaliteitsparameters van grondwaterproductielocaties benoemd.

4.1.1 Waterkwaliteit

Er is een positief verband gevonden tussen het TOC gehalte in het gedistribueerde water en het methaangehalte in ruwwater enerzijds en de nagroei-indicatoren anderzijds. Deze relaties komen ook naar voren in eerder BTO onderzoek (van der Wielen, 2011). De correlaties tussen het TOC gehalte van het reinwater en het ammoniumgehalte van het ruwwater enerzijds, en het KG22 of *Aeromonas* van het gedistribueerde water anderzijds zijn in deze studie minder significant dan in de eerdere BTO studie. Hieruit blijkt dat de significantie van de correlatie ondanks een redelijke groepsgrootte (respectievelijk 38 en 43) afhangt van de set van productielocaties die in de regressieanalyse betrokken zijn, of van de periode waaruit de data afkomstig zijn. Een andere verklaring voor de gevonden verschillen kan de factor tijd zijn aangezien voorliggende studie en het eerdere BTO onderzoek gebaseerd zijn op waterkwaliteitsdata uit verschillende perioden. Een deel van de variantie van de nagroei-indicatoren kan verklaard worden aan de hand van de betreffende ruwwaterkwaliteitsparameters. Echter, de afwezigheid van een eenduidige relatie tussen de ruwwaterkwaliteit en nagroei-indicatoren en de gevonden verschillen tussen voorliggende studie en het eerdere BTO onderzoek laten zien dat onderzoek naar de relatie tussen nagroei en andere parameters zoals zuivering ontwerp- en bedrijfsvoeringsgerelateerde variabelen noodzakelijk is.

4.1.2 Ontwerp zuivering

Er zijn verschillende verbanden gevonden tussen ontwerpeigenschappen en de waarden van nagroei-indicatoren.

Met betrekking tot het **type beluchting** geldt dat de productielocaties met een plaatbeluchting hogere nagroei-indicatorwaarden tonen, terwijl de productielocaties met cascade-, BOT- of compressorbeluchting vaker lagere nagroei-indicatorwaarden tonen. Voor de onderzochte productielocaties geldt in algemene zin dat de intensiteit van de beluchting zwaarder is naarmate het ruwwater meer ammonium en/of TOC bevat; deze positieve relatie is ook met methaan gevonden, al is die relatie minder sterk dan met de voornoemde parameters. Dit beeld is volgens verwachting, omdat het type beluchting bij het ontwerp

wordt afgestemd op de ruwwaterkwaliteit. Sproeibeluchting heeft een beperkte beluchtingseffectiviteit en vormt een uitzondering op het geschetste beeld omdat de sproeibeluchting ook bij ruwwaterklassen met hoge ammonium- en/of TOC-gehalten wordt toegepast. Op basis van een analyse van het type beluchting, de ruwwaterclassificatie op basis van ammonium en TOC, en de nagroei-indicatoren van het gedistribueerde water lijken productielocaties met sproeibeluchting vaker water te produceren met een lagere biologische stabiliteit. Ook is er een aanwijzing dat BOT beluchting gunstiger lijkt dan plaatbeluchting voor de productie van biologisch stabiel water. Mogelijk speelt de biomassa van methanotrofen hier een rol bij, gerichte metingen naar methanotrofen (met qPCR) geeft duidelijkheid of dat het geval is.

Op basis van een analyse van het genormaliseerd samengesteld nagroeigetal (*Aeromonas* en KG22) van het reinwater en de situatie met betrekking tot de **bovenwaterstand** lijkt een constante bovenwaterstand gunstiger voor de biologische stabiliteit dan een oplopende bovenwaterstand. De bovenwaterstand bij locaties met een constante bovenwaterstand is lager dan de gemiddelde bovenwaterstand bij locaties met een oplopende bovenwaterstand. Hierdoor is deze bevinding in overeenstemming met de bevinding met betrekking tot de bovenwatercontacttijd. Een oplopende bovenwaterstand kan daarnaast leiden tot een daling van de effectieve beluchting in het geval van sproeibeluchting.

Op basis van een analyse van de *Aeromonas*-aantallen en KG22 in het reine en gedistribueerde water en het **type filtratie** lijkt drukfiltratie mogelijk gunstiger voor de productie van biologisch stabiel dan gravitatiefiltratie.

4.1.3 Bedrijfsvoering

Er zijn verschillende verbanden gevonden tussen bedrijfsvoeringsgerelateerde eigenschappen en de waarden van nagroei-indicatoren.

Er is een positief verband gevonden tussen de **lucht-waterverhouding** (geaggregeerd op locatieniveau) en de waarde van nagroei-indicatoren in het gedistribueerde water. Dit vormt een aanwijzing dat een lage lucht-waterverhouding mogelijk gunstiger is voor de productie van biologisch stabiel water. Deze bevinding kan met twee mogelijke hypothesen uitgelegd worden: 1) vanwege de ruwwatersamenstelling is een hoge lucht-waterverhouding noodzakelijk en de verwijdering van chemische componenten zoals ijzer, mangaan en ammonium verloopt daardoor verondersteld effectief. Dat voorkomt dat deze stoffen doorslaan naar het drinkwater maar het kan leiden tot meer biomassavorming in het filter. De hogere biomassavorming in het filter kan leiden tot hogere waarden van nagroei in het gedistribueerde water. 2) Een hogere lucht-waterverhouding is (mede) verantwoordelijk voor een hogere nagroei doordat bijvoorbeeld AOC gevormd wordt in de beluchting of doordat componenten (anders dan zuurstof) die in de lucht aanwezig zijn ondanks luchtbehandeling aan het water worden overgedragen en aldaar tot nagroei kunnen leiden.

De analyse van het gecombineerd effect van de variabelen TOC en methaan in ruwwater en de eigenschappen van de beluchting (type, effectief beluchtingsgetal en lucht-waterverhouding) op de waarden van nagroei-indicatoren onderschrijft de hierboven beschreven relaties maar nader onderzoek is nodig om te bepalen welke van deze parameters het belangrijkste zijn in die relaties.

Er is een positief verband gevonden tussen de **bovenwatercontacttijd** van de filters en de waarde van het KG22 in het gedistribueerde water. Dit vormt een aanwijzing dat een lage bovenwatercontacttijd mogelijk gunstiger is voor de productie van biologisch stabiel water. Mogelijk wordt deze observatie beïnvloed door het verband met de contacttijd van water en

lucht in geval van sproeibeluchting. Deze contacttijd is lager bij een hogere bovenwaterstand en een hogere bovenwatercontacttijd. Een andere mogelijke oorzaak is dat er in het bovenwater processen optreden die een nadelig effect hebben op de biologische stabiliteit van het water. Overigens is er een positief verband gevonden tussen de bovenwatercontacttijd en de verwijdering van ijzer. Daarbij is niet onderzocht van welk type ontijzering sprake is (homogeen, heterogeen of biologisch) en wat de mogelijke relatie met het ontijzeringsmechanisme is.

Er is een positief verband gevonden tussen het **spoelwaterpercentage** en het genormaliseerde KG22 van het gedistribueerde water; een vergelijkbaar verband is gevonden tussen het genormaliseerd spoelluchtgetal en het genormaliseerd samengesteld nagroei getal. Dit vormt een aanwijzing dat een lager spoelwater- of spoelluchtpercentage een positief effect heeft op de biologische stabiliteit van het water. In beide gevallen zou eerder een negatief verband verwacht worden: naarmate het spoelwater- of spoellucht percentage toeneemt is de spoeling verondersteld beter, wordt gewenst materiaal effectief uit het filter verwijderd en komt dat ten goede aan de waterkwaliteit. Een mogelijke verklaring voor het positieve verband is dat de productielocaties die op basis van de ruwwaterkwaliteit een omvangrijk spoelwaterprogramma nodig hebben, ondanks dat omvangrijke spoelprogramma, water produceren met een hogere nagroei. Een andere verklaring is mogelijk dat het spoelwaterprogramma "te" omvangrijk is, waardoor bijvoorbeeld bepaald materiaal van chemische of biologische samenstelling van het filtermedium wordt losgemaakt. Dat zou negatief kunnen werken in het geval dat het losgemaakte materiaal niet effectief uit het filter wordt verwijderd, wanneer het materiaal wel effectief wordt verwijderd maar het losmaken en de verwijdering ongewenst is, of dat door het spoelen geen goede opbouw van bacteriële biomassa plaatsvindt over het filter.

Met betrekking tot de **variabiliteit** van de bedrijfsvoering is geen relatie gevonden tussen de waarden van nagroei-indicatoren en de wijze van inbedrijfname van een filter (direct of geleidelijk), de schakelfrequentie van filters, de schakelfrequentie van pompputten en de mate van variatie van productiedebieten. Wel is een positief verband gevonden tussen het aantal parallelle filterunits en de nagroei-indicatoren: alle productielocaties met tien of meer parallelle units (6 van de 43 locaties hebben tien units of meer) tonen lagere aantallen van de nagroei-indicatoren in het gedistribueerde water. Productielocaties met een groter aantal parallelle units hebben mogelijk minder last van discontinuïteiten in de bedrijfsvoering omdat effecten opgevangen worden door een groot aantal filterunits. Deze productielocaties met een groot aantal parallelle units hebben in algemene zin een lage filtratiesnelheid. De filtratiesnelheid heeft mogelijk een effect, echter het is niet de verklarende oorzaak van het positieve effect op de nagroei-indicatoren, omdat ook is aangetoond dat verschillende productielocaties met een lage filtratiesnelheid hoge nagroei-indicatorwaarden tonen (paragraaf 3.6.3).

Gebaseerd op een mogelijk hypothese ten aanzien van de relatie tussen **biologische ontijzering** en nagroei is een analyse uitgevoerd van de gecombineerde gegevens van het ijzergehalte en de pH van het ruwwater, het effectief beluchtingsgetal en een nagroei-indicator. Hieruit blijkt dat productielocaties met een hoge pH en een laag ijzergehalte in het ruwwater een laag effectief beluchtingsgetal en vaak lage waarden van nagroei-indicatoren hebben. Bij deze locaties met een laag ijzergehalte en een hoge pH is ontijzering waarschijnlijk niet biologisch van aard. Productielocaties met de hoogste nagroei-indicatoren hebben een hoger ijzergehalte, een lagere pH en een hoger effectief beluchtingsgetal. Bij deze locaties met een hoog ijzergehalte en een lage pH kan de ontijzering biologisch van aard zijn. Op basis van deze bevinding kan de hypothese gesteld worden dat de aanwezigheid van biologische ontijzering mogelijk bijdraagt aan de nagroei. De gevonden

observaties verklaren niet alle variantie van de nagroei-indicatoren, want er zijn ook productielocaties die mogelijk (gedeeltelijk) biologisch ontijzeren en waarbij de nagroei-indicatoren laag zijn. Uit dezelfde analyse met de lucht-waterverhouding in plaats van het effectief beluchtingsgetal volgt een vergelijkbaar beeld. Op basis van het vergelijken van de locaties met een hoog ijzergehalte en een lage pH kan aanvullend worden afgeleid dat de locaties die op basis van de ruwwaterkwaliteit mogelijk geschikt zijn voor biologische ontijzering én ook een lage lucht-waterratio hebben – dat is gunstig voor de biologische ontijzering – beter presteren dan de locaties met vergelijkbare ruwwaterkwaliteit (ijzer en pH) en een hogere lucht-water ratio. Dit sluit aan bij het positieve verband tussen de nagroei-indicatoren en de lucht-water verhouding.

Uit een beperkt aantal meervoudige analyses van een waterkwaliteitsparameter, een bedrijfsvoeringsparameter (met continue waarde) en een ontwerp- of **bedrijfsvoeringsparameter met binaire waarde** komt naar voren dat de volgende eigenschappen gunstig zijn voor de biologische stabiliteit: i) er is sprake van filterschoonmaak, ii) er is eerste filtraatafvoer, iii) er is geen sprake van spoelwaterhergebruik, iv) er is sprake van een continue bedrijfsvoering van de beluchting of v) er is geen sprake van een continue bedrijfsvoering van de filters. De bevindingen met betrekking tot filterschoonmaak, eerste filtraatafvoer, spoelwaterhergebruik en de continuïteit van de beluchting kunnen aansluiting vinden bij de verwachtingen t.a.v. het vergelijken van de prestaties van een productielocatie die deze eigenschappen wel of niet bevat. Hierbij is niet gezegd dat een productielocatie waarbij bijvoorbeeld geen sprake is van filterschoonmaak slecht presteert, maar ten opzichte van een locatie waarbij de filters wel worden schoongemaakt mag verwacht worden dat de kans dat de filterschoongemaakte locatie beter presteert groot is. De bevinding met betrekking tot de continuïteit van de filters is onverwacht. Op basis hiervan kan een hypothese opgesteld dat de biologische stabiliteit verbetert wanneer sprake is van filterschoonmaak, eerste filtraatafvoer of geen sprake van spoelwaterhergebruik.

4.2 Relaties met chemische parameters

In deze paragraaf zijn de gevonden verbanden tussen de waarden van chemische en fysische waterkwaliteit parameters en de ontwerp- en bedrijfsvoeringsgerelateerde eigenschappen en overige waterkwaliteitsparameters van grondwaterproductielocaties beschreven. In voorliggende studie zijn de relaties tussen het ontwerp, de bedrijfsvoering en de chemische waterkwaliteit onderzocht omdat verondersteld is dat de aanwezigheid van zulke opgeloste en deeltjesgebonden stoffen een negatieve invloed kunnen hebben op de biologische stabiliteit van reinwater en bijdragen aan nagroei in het leidingnet. Met de toegepaste dataset zijn geen (negatieve) verbanden gevonden tussen de waarden van de nagroei-indicatoren in reinwater of gedistribueerd water en de waarden van het ijzergehalte of de troebelingsgraad in reinwater of gedistribueerd water.

De verlaging van het samengesteld chemisch getal (conversie van ijzer, mangaan, ammonium) heeft een positieve relatie met de filterbedcontacttijd en het spoelwaterpercentage, en een negatieve relatie met de filtratiesnelheid. Vergelijkbare verbanden zijn ook gevonden voor de conversie van de afzonderlijke componenten (ijzer, mangaan, ammonium) en TOC. De hogere ijzergehalten in reinwater (tot 0,05 mg/L) komen voor bij de productielocaties die een laag ijzergehalte in het ruwwater hebben. De contacttijd en de filtratiesnelheid zijn aan elkaar gecorreleerd maar er is geen sprake van een een-op-een-relatie. Het gevonden verband met de contacttijd en de filtratiesnelheid is volgens verwachting: een voldoende lange contacttijd en beperking van de afschuifspanning zijn gunstig voor de effectieve omzetting en verwijdering van deze chemische parameters. Het

gevonden verband met het spoelwaterpercentage is volgens verwachting: een effectieve spoeling is gunstig voor de verwijdering van de betreffende parameters.

5 Conclusies en aanbevelingen

In dit hoofdstuk zijn de conclusies beschreven met betrekking tot het opstellen en het toepassen van de database, is beschreven in hoeverre het onderzoek voldoet aan de projectdoelstellingen en zijn de belangrijkste inhoudelijke conclusies beschreven met betrekking tot gevonden aanwijzingen voor optimalisatie en verbetering van grondwaterzuiveringen. Tevens zijn de aanbevelingen beschreven die uit het onderzoek volgen.

5.1 Conclusies

Ten aanzien van de **overall projectdoelstelling** geldt de volgende conclusie:

Er is geen enkelvoudige relatie gevonden die de variantie volledig en nauwkeurig bepaalt tussen de ruwwaterkwaliteit, de ontwerp- of bedrijfsvoeringsgerelateerde eigenschappen enerzijds en biologische nagroei-indicatoren anderzijds. Wel zijn verschillende enkelvoudige relaties gevonden die afzonderlijk of gezamenlijk verklarend kunnen zijn voor een deel van de waargenomen variantie van de nagroei-indicatoren. Daarnaast is een beperkt aantal meervoudige analyses (met drie of vier parameters) gericht op “handmatige” gekozen hypothesen op kwalitatieve wijze onderzocht. Deze methode is tijdrovend, maar lijkt geschikt om op hoofdlijnen inzicht te verkrijgen in de samenhang tussen drie of vier parameters.

Met betrekking tot de **database en data-analyse** gelden de volgende conclusies:

- De dataset laat zich als volgt kenmerken: (i) er is heterogene variëteit in de attributen (verschillende procesconfiguraties), (ii) er zijn hiërarchische eigenschappen (eigenschappen kunnen slechts een waarde hebben als een voorgaande eigenschap een waarde heeft), (iii) de matrix is vanwege verschillende redenen ‘dun’ gevuld (de matrix is *sparse*), (iv) er zijn relatief veel bekende eigenschappen per zuivering in relatie tot het aantal zuiveringen. Door de grote variëteit aan informatie en een relatief klein aantal onafhankelijke observaties wordt een data-analyse bemoeilijkt. Gezien deze kenmerken is het trainen van complexere, *machine-learning* regressiemodellen niet de aangewezen analyseroute. Om een alternatieve aanpak te volgen waarbij enkelvoudige relaties en hypothesegedreven meervoudige relaties onderzocht zijn, is een *machine learning toolbox* verder uitgewerkt in de programmeertaal Python. Deze toolbox heeft de volgende functionaliteit: importeren data uit bronbestanden, afgeleide eigenschappen berekenen en toevoegen, en voorbewerking van de dataset (o.a. verwijdering van eigenschappen met incomplete informatie), regressieanalyse en datavisualisatie.
- Geconcludeerd wordt dat de database in de huidige vorm een bruikbare bron van informatie is die ingezet kan worden bij onderzoek naar de biologische stabiliteit van water, als naslagwerk, om de prestatie en eigenschappen van overige zuiveringen te vergelijken, en naar verwachting ook bij onderzoek, ontwerp en optimalisatie van grondwaterzuiveringen in bredere zin.

Met betrekking tot de **inhoudelijke relaties** tussen nagroei-indicatoren, waterkwaliteit, en ontwerp- en bedrijfsvoeringsgerelateerde eigenschappen van de zuivering zijn een aantal conclusies opgesteld. De relaties tussen een groot aantal eigenschappen is onderzocht. Bij de hieronder genoemde parameters is een relatie met de nagroei-indicatoren KG22 en *Aeromonas* gevonden, wat een eerste aanwijzing vormt dat deze parameters de mate van nagroei mogelijk beïnvloeden. Deze aanwijzingen zijn gebruikt om hypothesen op te stellen

(zie Hoofdstuk 4), waarmee de invloed van deze parameters op de nagroei-indicatoren gerichter kan worden onderzocht in vervolgonderzoek.

- **Waterkwaliteit.** het TOC-gehalte in het gedistribueerde water en het methaangehalte in ruwwater tonen een positieve relatie met de nagroei-indicatoren.
- **Ontwerp zuivering:**
 - Type beluchting. Sproeibeluchting lijkt minder gunstig voor de productie van biologisch stabiel water. BOT-beluchting lijkt gunstiger dan plaatbeluchting voor de productie van biologisch stabiel water.
 - Bovenwaterstand. Een constante bovenwaterstand lijkt gunstiger dan een oplopende bovenwaterstand voor de productie van biologisch stabiel water.
 - Type filtratie. Drukfiltratie lijkt gunstiger dan gravitatiefiltratie voor de productie van biologisch stabiel water.
- **Bedrijfsvoering zuivering:**
 - Lucht-waterverhouding. Er is een positief verband gevonden tussen de lucht-waterverhouding en de waarde van nagroei-indicatoren in het gedistribueerde water.
 - Bovenwatercontacttijd. Er is een positief verband gevonden tussen de bovenwatercontacttijd van de filters en de waarde van het KG22 in het gedistribueerde water.
 - Filterspoeling. Er is een positief verband gevonden tussen het spoelwaterpercentage en het genormaliseerde KG22 van het gedistribueerde water. Een vergelijkbaar verband is gevonden tussen het genormaliseerd spoelluchtgetal en het genormaliseerd samengesteld nagroei-getal.
 - Variabiliteit. Er is een positief verband gevonden tussen het aantal parallelle filterunits (een maat voor het uitdempen van variaties in de zuivering) en de nagroei-indicatoren. Er is geen relatie gevonden tussen de waarden van nagroei-indicatoren en de variabiliteit van de bedrijfsvoering gebaseerd op de wijze van inbedrijfname van een filter (direct of geleidelijk), de schakelfrequentie van filters, de schakelfrequentie van pompputten en de mate van variatie van productiedebieten.
 - Biologische ontijzering. Er zijn aanwijzingen gevonden voor mogelijk negatief effect van biologische ontijzering op de biologische stabiliteit van drinkwater. Voor de locaties waar de ontijzering mogelijk biologisch van aard is, geldt dat de locaties met een lage lucht-waterverhouding beter presteren.
 - Optimale filtratiesnelheid. Er is geen verband gevonden tussen de filtratiesnelheid en nagroei-indicatoren, wat duidt op het afwezig zijn van een optimale filtratiesnelheid.
 - Bedrijfsvoeringsparameter met binaire waarde. Filterschoonmaak, eerste filtraatafvoer, en de afwezigheid van spoelwaterhergebruik lijken gunstig voor het KG22 in het gedistribueerde water.

Relaties chemische en fysische parameters:

De verlaging van het samengesteld chemisch getal (conversie ijzer, mangaan, ammonium) heeft een positieve relatie met de filterbedcontacttijd en het spoelwaterpercentage, en een negatieve relatie met de filtratiesnelheid. Vergelijkbare verbanden zijn ook gevonden voor de conversie van de afzonderlijke componenten (ijzer, mangaan, ammonium) en TOC. De hogere ijzergehalten in reinwater (tot 0,05 mg/L) komen voor bij de productielocaties die een laag ijzergehalte in het ruwwater hebben.

5.2 Aanbevelingen

De database is in de huidige vorm een bruikbare bron van informatie en in te zetten bij onderzoek naar de biologische stabiliteit van water, als naslagwerk, en bij onderzoek, ontwerp en optimalisatie van grondwaterzuiveringen in bredere zin⁷. De kwaliteit van de database kan op verschillende manieren verder verbeterd worden:

- De kwaliteit van de informatie in de database kan verder verbeterd worden met behulp van afstemming met de drinkwaterbedrijven.
- De database kan verder gecombineerd worden door ontbrekende relevante eigenschappen te achterhalen in overleg met de drinkwaterbedrijven.
- De database kan uitgebreid worden met nieuwe productielocaties.
- De database kan uitgebreid worden met nieuwe (afgeleide) parameters, zoals de filterbelasting (vracht per m² of m³ filterbed), het fosfaatgehalte of het AOC gehalte.

Uitbreiding van de database, door toevoeging van productielocaties dan wel door toevoeging van ontbrekende gegevens, is aanbevelingswaardig (i) op basis van de bevinding dat de significantie van gevonden correlaties in voorliggend en eerder BTO onderzoek ondanks een redelijke groepsgrootte (respectievelijk 43 en 38) afhangt van de set van productielocaties die in de studie betrokken zijn of van de periode waaruit de data afkomstig zijn, en (ii) omdat de mogelijkheden voor het uitvoeren van meervoudige analyses daarmee toenemen.

In het kader van de verdere verbetering van de biologische stabiliteit kan vervolgonderzoek zich richten op de volgende onderdelen:

- Gericht onderzoek naar de vastgestelde hypothesen die zijn opgesteld aan de hand van de vastgestelde relaties tussen ruwwaterkwaliteit, ontwerp zuivering, bedrijfsvoering zuivering en de nagroei-indicatoren, zoals geformuleerd in Hoofdstuk 4. Dit onderzoek kan op laboratorium-, proefinstallatie- en praktijkschaal plaatsvinden.
- De eigenschappen van productielocaties die uit verschillende zuiveringsunits (in serie en/of parallel) zijn geaggregeerd tot locatieniveau. Om het effect van de onzekerheid die met deze aggregatie gepaard gaat te elimineren kan de correlatieanalyses uitgevoerd worden voor een selectie van de productielocaties, namelijk diegene die uit één straat of één unit bestaan.
- Bij verschillende productielocaties is sprake van een verstoring effect door de aanwezigheid van zuiveringsstappen anders dan beluchting en filtratie of door bijmenging van water van een andere locatie. Om het effect van de onzekerheid die met deze verstoring gepaard gaat te elimineren kunnen de correlatieanalyses uitgevoerd worden voor een selectie van de productielocaties, namelijk diegene waarbij geen sprake is van een verstoring.
- In voorliggend onderzoek is een beperkt aantal hypothesen op "handmatige" wijze vastgesteld dan wel arbitrair geselecteerd, en deze zijn middels een meervoudige analyse (met drie of vier parameters) op kwalitatieve manier getoetst. Deze methode is tijdrovend en derhalve is het aantal geformuleerde en geteste hypothesen in voorliggend onderzoek niet uitputtend. Een statistisch onderbouwde meervoudige regressieanalyse kan handmatig worden uitgevoerd met SPSS, maar een dergelijke studie zal relatief veel tijd kosten. Softwarematige automatisering van een uitputtende meervoudige analyse, door het testen van alle mogelijke of gewenste combinaties van drie of vier parameters, is mogelijk met behulp van logistische regressie en parameter optimalisatie routine. Vanwege de eigenschappen van de database (*sparse* matrix en grote verhouding tussen

⁷ In een BTO onderzoeksvoorstel over ontijzering van de Themagroep 'Drinkwatertechnologie van de Toekomst' is voorgesteld om de database te gebruiken. Het voorstel is door het CO goedgekeurd en zal naar verwachting in 2016 starten.

- eigenschappen en aantal locaties) is een belangrijke randvoorwaarde bij die analyse dat het aantal parameters in de regressie gemaximeerd wordt tot drie of vier. De betrouwbaarheid van de regressie zou getoetst kunnen worden door validatie van het model na toevoeging van een nieuwe locatie. *Principal component analysis* met een beperking van aantal attributen is een tweede mogelijkheid om meervoudige regressie uit te voeren.
- Toepassing van de database door (nieuwe) grondwater productielocaties met hogere waarden van de nalgroei indicatoren "handmatig" te vergelijken met de prestaties, het ontwerp en de bedrijfsvoering van vergelijkbare zuiveringen.

6 Referenties

6.1 Referenties

- Bereschenko, L.A., 2014, Inventarisatie nagroeiproblematiek Oasen en oorzaak lagere biologische stabiliteit zs De Laak, BTO 2014.206(s), KWR Watercycle Research Institute, Nieuwegein.
- Domingos, P., 2012, A Few Useful Things to Know About Machine Learning, Communications of the ACM, vol. 55, pp 78-86.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, A., Thirion, B.V., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., 2011, Scikit-learn: Machine learning in Python, Journal of Machine Learning Research, vol. 12, pp 2825-2830.
- Reijnen, G.K., 1994, Behandeling van methaanhoudend grondwater, Kiwa mededeling 123.
- Van der Helm, A.W.C., 1998, Modelling van intensieve gasuitwisselingssystemen, MSc thesis, TU Delft.
- Van der Kooij, D. & Veenendaal, H., 2012, Bepaling van de biofilmvormende eigenschappen van drinkwater met een continue biofilmmonitor (CBM), BTO 2011.050, KWR Watercycle Research Institute, Nieuwegein.
- Van der Kooij, D. & Veenendaal, H., 2013, Bepaling van de concentratie van de biomassa in drinkwater met behulp van hemoflow, BTO 2013.044, KWR Watercycle Research Institute, Nieuwegein.
- Van der Kooij, D. & Veenendaal, H., 2014, Bepaling van de biomassaproductiepotentie (BPP) van drinkwater, BTO 2014.038, KWR Watercycle Research Institute, Nieuwegein
- van der Kooij, D. & Bereschenko, L.A., 2015, De invloed van ijzer op de biomassa-accumulatie in de continue biofoulingmonitor (CBM). BTO 2015.216(s), KWR Watercycle Research Institute, Nieuwegein.
- Van der Wielen, P. & Van der Kooij, D., 2011, Inventarisatie van *Aeromonas* en koloniegetal 22°C in drinkwater en relaties met fysisch/chemische parameters, BTO 2011.001, KWR Watercycle Research Institute, Nieuwegein.
- Van Leerdam, R., 2012, Vergelijking tussen een aantal grondwaterproductiebedrijven op biologische stabiliteitsparameters in relatie tot het zuiveringsproces, BTO 2012.223(s), KWR Watercycle Research Institute, Nieuwegein.
- Vonk, E. & Vries, D., 2015, Datamining voor assetmanagement - inventarisatie en voorbeelden uit de watersector, BTO 2015.077, KWR Watercycle Research Institute, Nieuwegein.

Bijlage I Lijst ontwerp en bedrijfsvoeringsgerelateerde eigenschappen

In deze bijlage is in meer detail beschreven uit welke ontwerp en bedrijfsvoeringsparameter het format is opgebouwd. In het feitelijke format staan meer attributen.

Per eigenschap is toegelicht of en op welke wijze deze is toegepast in de correlatie analyse. Tevens is toegelicht op welke wijze aggregatie van unit-niveau tot locatie niveau is uitgevoerd. In de kolom "focus correlatie analyse" is aangegeven op welke eigenschappen op voorhand het meest gelet is. In principe zijn in het *heatmap* correlatie overzicht alle mogelijke 1-op-1 correlaties getest, zodat eventuele sterke correlaties met eigenschappen waar op voorhand niet op gefocust is toch zullen verschijnen. De laatste kolom geeft voor bepaalde eigenschappen het aantal punten in de database weer.

Parameter	Toelichting gebruik	Focus bij correlatie analyse?	Aantal datapunten/ aantal locaties met deze eigenschap
ALGEMENE LOCATIE PARAMETERS			
Naam productielocatie	Zichtbaar in plots met behulp van de cursor.	Als indicator	
Naam waterbedrijf	Indirect bekend via de naam van de productielocatie.	-	
Het bouwjaar van de locatie (oorspronkelijk en na eventuele renovatie)	Omgerekend naar de leeftijd van de productielocatie.	ja	LT1: allen LT2: 17
Ontwerpcapaciteit en gemiddelde capaciteit	Omrekenen naar de benuttingsgraad (ratio gemiddeld : ontwerp)	ja	allen
Maximale jaarcapaciteit	Niet gebruikt, geeft ten opzichte van de ontwerp of de gemiddelde capaciteit voor veel productielocaties geen nieuwe informatie.	-	allen
Volstroom of deelstroom behandeling	Niet gebruikt, niet onderscheidend (alle locaties 'volstroom').	-	allen
Disturbance (aanwezigheid van een extra zuiveringsstap naar beluchting en zandfiltratie, of bijmengen van ander reinwater)	Eventueel te gebruiken als verklarende informatie bij gevonden verband.	- Indirect	14 x TRUE
Spoelwaterhergebruik		ja	11 x TRUE
Aantal beluchtingsstappen in serie		ja	allen
Maximaal en gemiddeld waterdebiet voor totale beluchtingssectie	Omrekenen naar benuttingsgraad (ratio maximaal : gemiddeld) per beluchtingsstap. Niet op het niveau van beluchtingsstap beschouwen maar aggregeren op locatieniveau:	-	
PB_LOC_BS_debiet_GemMax_Ratio_TOT	In geval dat de locatie meerdere beluchtingsstappen heeft, zijn de afzonderlijke benuttingsgraden als volgt geaggregeerd tot een gemiddelde op locatieniveau: Gewogen gemiddelde op basis van het gemiddeld debiet (voor zowel serie als parallel schakeling). Volgorde: binnen 1 tak eerst serie middelen (als van toepassing), dan parallel, en dan weer serie als parallel eventueel weer samenkomt.	ja	40
Minimaal waterdebiet voor totale beluchtingssectie	In deze analyse nog niet geselecteerd, mede wegens relatief beperkt bekende data.	-	28
Maximaal en gemiddeld waterdebiet per beluchtingsunit	Niet op het niveau van beluchtingsstap beschouwen maar aggregeren op locatieniveau:	-	
PB_LOC_BS_debiet_Max_perBS PB_LOC_BS_debiet_Gem_perBS	In geval dat de locatie meerdere beluchtingsstappen heeft, zijn de afzonderlijke (gem en max) debieten geaggregeerd tot een gemiddelde op locatieniveau: Gewogen gemiddelde (voor zowel serie als parallel schakeling). Volgorde middelen (indien van toepassing): binnen 1 tak eerst serie middelen, dan parallel, en dan weer serie als parallel eventueel weer samenkomt. Bv Y-type zuivering: eerst parallel middelen, dan gemiddelde van de 2 stappen in serie die ontstaan. Rekening houden met totaal debiet per straat.	ja	40

Waterdebiet door de beluchtingssectie (%) ... _DebietBeluchting%	gebruikt in berekeningen	-	allen
Type beluchting	Gebruikt als indicator in overzichtsplaatje Gebruikt in correlatie analyse middels clustering	- ja	allen
PARAMETERS BELUCHTING			
Effectief beluchtingsgetal PB_LOC_BS_isEFF	Inschatting effectiviteit beluchting om type beluchting te kwantificeren. Daarmee worden verschillende beluchtingsstappen onderling beter vergelijkbaar. In het effectief beluchtingsgetal worden zo goed als mogelijk de eigenschappen van de beluchting meegenomen. In aparte paragraaf berekeningswijze beschrijven.	ja	allen
Eigenschappen cascade: - Aantal cascades - Aantal valstappen - Valhoogte per stap - Breedte bak - Hoogte bak - Diepte bak	Gegevens gebruikt als input voor volgende berekeningen: - Debiet per beluchtingsunit - Totale valhoogte water - Contacttijd	ja ja ja	14xCAS
Verdeling water cascade in stralen? PB_LOC_BS1_unit1_Cas_isVerdelingWater	Wordt het water in de cascade in separate stralen verdeeld. Nog niet meegenomen in correlatie analyse vanwege focus en beperkt aantal data. Eventueel gebruiken voor aanvullende verklaring.	-	
Aantal cascades in bedrijf constant? PB_LOC_BS1_unit1_Cas_isConstant	Nauwelijks onderscheidend. Voor alle beluchtingsstappen aggregeren tot locatie niveau en in twee groepen (constant, niet constant) vergelijken	- ja	13
Aantal sproeistappen in serie binnen deze sproeibeluchting PB_LOC_BS1_unit1_SproeiStapSerie	Niet gebruikt omdat deze parameter niet onderscheidend is (meestal 1 sproeistap in serie binnen de beluchtingsstap)	-	38xsproei
Is er een sproeikamer? PB_LOC_BS1_unit1_isSproei1Kamer PB_LOC_BS1_unit1_isSproeimeerKamer	Niet gebruikt omdat deze parameter niet onderscheidend is (meestal geen separate sproeikamer) ALS "2 sproeistappen": Eerst sproeien in separate kamer en daarna sproeien direct boven filter? Niet gebruikt, komt nauwelijks voor in database.	- -	
Type sproeier (Dresden, Limburg, WMN, Gouds, KSH, sproei-arm, Overig) PB_LOC_BS1_unit1_SproeiStap_Label	In deze analyse nog geen focus, eventueel te gebruiken als verklarende informatie bij gevonden verband.	-	
Opwaarts of neerwaarts sproeien? PB_LOC_BS1_unit1_SproeiStap_isOpwaarts	Niet gebruikt omdat deze parameter nauwelijks onderscheidend is (in 2 van de 20 gevallen opwaarts)	-	
Valhoogte water: - Bodem tot kamer - Sproeier tot bovenkant filter - Sproeier tot bodem	In de meeste gevallen is de afstand sproeier tot bovenkant filterbed bekend. Deze valhoogte is gebruikt in berekening contacttijd en het effectiviteitsgetal beluchting	ja	33
Aantal nozzles per sproei unit PB_LOC_BS1_unit1_aantalNozzles	Omrekenen naar aantal nozzles per m2 filter, eventueel gebruik binnen het cluster 'sproeiers'	weinig data	9
Aantal nozzles in bedrijf constant? PB_LOC_BS1_unit1_Nozzle_isConst	Aantal nozzles in bedrijf constant of variabel. Nauwelijks onderscheidend (meestal constant). Voor alle beluchtingsstappen aggregeren tot locatie niveau en in twee groepen (constant, niet constant) vergelijken	- ja	12

Aantal BOT per filter en totaal aantal BOTs PB_LOC_BS1_unit1_AantalBOTpFilt PB_LOC_BS1_unit1_AantalBOT_TOT	Gebruikt in berekening debiet per beluchtingsunit. In deze analyse nog geen focus, eventueel te gebruiken als verklarende informatie bij gevonden verband.	-	16xBOT
Type pakking (Pallring, Rashing, Overig)	In deze analyse nog geen focus, eventueel te gebruiken als verklarende informatie bij gevonden verband.	-	
Materiaal pakking PB_LOC_BS1_unit1_FiltPakVul_isKunststof	Niet gebruikt in analyse, parameter niet onderscheidend (meestal kunststof pakking).	-	
Leeftijd pakking PB_LOC_BS1_unit1_FiltPak_LT	Niet gebruikt in analyse, te weinig data.	-	
Dimensie BOT (bedhoogte en oppervlak) - PB_LOC_BS1_unit1_FiltPak_BedHoogte - PB_LOC_BS1_unit1_BOT_Oppervlak	Gebruikt voor berekening contacttijd	ja	13
Tegenstroom of meestroom beluchting in de BOT? PB_LOC_BS1_unit1_FiltPak_isTegenStr	In deze analyse nog geen focus, eventueel te gebruiken als verklarende informatie bij gevonden verband.	-	
Aantal BOTs in bedrijf constant? PB_LOC_BS1_unit1_FiltPak_isConst	Niet op BOT niveau analyseren Voor alle beluchtingsstappen aggregeren tot locatie niveau en in twee groepen (constant, niet constant) vergelijken	- ja	7
PB_LOC_BS1_unit1_FiltPak_isTerugSpoel	In deze analyse nog geen focus, eventueel te gebruiken als verklarende informatie bij gevonden verband.	-	
Aantal plaatbeluchters PB_LOC_BS1_unit1_Plaat_Aantal	Gebruikt in berekening debiet per beluchtingsunit. In deze analyse nog geen focus, eventueel te gebruiken als verklarende informatie bij gevonden verband.	-	5xplaat
Dimensies plaatbeluchting: lengte en breedte (oppervlak PB_LOC_BS1_unit1_Plaat_Lengte PB_LOC_BS1_unit1_Plaat_Breedte	Gebruik voor berekening contacttijd	ja	5
Dikte waterlaag PB_LOC_BS1_unit1_Plaat_DikteWater	Gebruik voor berekening contacttijd	ja	5
PB_LOC_BS1_unit1_Plaat_Gatpm2	In deze analyse nog geen focus, eventueel te gebruiken als verklarende informatie bij gevonden verband.	-	
PB_LOC_BS1_unit1_Plaat_Gat_diameter	In deze analyse nog geen focus, eventueel te gebruiken als verklarende informatie bij gevonden verband.	-	
Aantal plaatbeluchters in bedrijf constant? PB_LOC_BS1_unit1_Plaat_isConst	Niet op plaat beluchter niveau analyseren Voor alle beluchtingsstappen aggregeren tot locatie niveau en in twee groepen (constant, niet constant) vergelijken	- ja	4xFALSE 0xTRUE
Aantal compressor of venturi units. PB_LOC_BS1_unit1_Comp_unitperFilt PB_LOC_BS1_unit1_Comp_AantalABS	Gebruikt in berekening debiet per beluchtingsunit. In deze analyse nog geen focus, eventueel te gebruiken als verklarende informatie bij gevonden verband.	-	11xcomp
Aantal compressor of venture beluchters in bedrijf constant? PB_LOC_BS1_unit1_Comp_AantalisConst	Niet op niveau compressor/venturi analyseren Voor alle beluchtingsstappen aggregeren tot locatie niveau en in twee groepen (constant, niet constant) vergelijken	- ja	7
Is ontluchting aanwezig? PB_LOC_BS1_unit1_Comp_Ontluchting	Niet gebruikt in analyse, parameter niet onderscheidend (meestal ontluchting aanwezig).	-	
PB_LOC_BS1_unit1_Comp_Contacttijd	Voor elke situatie dezelfde contacttijd aangenomen. Orde grootte inschatting gemaakt op basis van een debiet, leiding diameter en afstand van injectie punt tot aan filter:	ja	

	aanname contacttijd gelijk aan 1 minuut.		
Zuurstof dosering	Geen locaties met zuurstofdosering in de database	-	-
Deep well aeration	Geen locaties met <i>deep well aeration</i> in de database	-	-
Beluchtingstappen locatie constant in bedrijf?	Voor alle beluchtingsstappen aggregeren tot locatie niveau en in twee groepen (constant, niet constant) vergelijken. Aggregeren: in de meeste gevallen is de waarde voor alle beluchtingsstappen van een locatie gelijk. Wanneer de waarde verschillend is, dan is de waarde van de locatie gelijk aan de meerderheid van de afzonderlijke waarden. In het geval dat er sprake is van 1 maal TRUE en 1 maal FALSE is de waarde van de locatie gelijk aan FALSE (d.w.z. de beluchting is niet constant in bedrijf).	ja	allen (24xTRUE)
Geforceerde beluchting?	Aggregeren: in de meeste gevallen is de waarde voor alle beluchtingsstappen van een locatie gelijk. Wanneer er sprake is van verschillende waarden is de waarde van de locatie gelijk aan TRUE omdat een deel van het water geforceerd wordt belucht/het water in een deel van de beluchtingsstappen geforceerd wordt belucht. Het effect van een mix van geforceerd en niet geforceerd beluchten komt tot uitdrukking bij de geaggregeerde lucht-water verhouding. In de meeste gevallen wordt altijd wel een deel van het water geforceerd belucht. In slechts 2 van de 43 gevallen wordt het water helemaal niet geforceerd belucht.	ja	43
Het aantal beluchtingscompressoren PB_LOC_BS1_unit1_isForc_AantalpUnit PB_LOC_BS1_unit1_isForc_AantalTOT	In deze analyse nog geen focus, eventueel te gebruiken als verklarende informatie bij gevonden verband.	-	
Is het luchtdebiet tijdens geforceerde beluchting constant? PB_LOC_BS1_unit1_isForc_isConst	Niet gebruikt, nauwelijks onderscheidend. Wel geaggregeerd op locatieniveau.	-	
Luchtdebiet per beluchting PB_LOC_BS1_unit1_LuchtDebiet_TOT PB_LOC_BS1_unit1_LuchtDebiet_pUnit PB_LOC_BS1_unit1_LuchtWater_Ratio	Gebruiken om de lucht-water verhouding te berekenen Lucht-water verhouding (RQ) aggregeren tot locatie niveau: sommen van de RQ bij beluchting in serie en debiet gewogen middelen bij parallelle beluchting.	- ja	 30
Wordt het luchtdebiet aan de hand van het waterdebiet geregeld? PB_LOC_BS1_unit1_isLuchtRegelWater	In 3 van 43 gevallen is er sprake van een regeling. Geen focus, te weinig data	-	3
Wordt de lucht behandeld? PB_LOC_BS1_unit1_isLuchtBehandeld	In de meeste gevallen is sprake van een combinatie van een grofstof filter, een fijnstof filter en/of een absoluutfilter. In 1 van de 43 locaties wordt de lucht niet behandeld (Goor). In 3 gevallen (Herikerberg en Onnen, Weerselo) is sprake van minimale behandeling (vliegengaas). Deze locaties hebben geen opvallend hoge nagroei parameters.	ja	

wordt de beluchting frequent schoongemaakt, zo ja hoe vaak? PB_LOC_BS1_unit1_isBeluchtingSchoonmaak	Aggregeren tot locatieniveau. Als een van de beluchtingsstappen wordt schoongemaakt dan is de waarde voor de locatie TRUE. Weinig onderscheidend.	-	39xTRUE 2xFALSE
PB_LOC_BS1_unit1_isBeluchtingSchoonmaakFrequentie	Frequentie eventueel middelen.	ja	37
is de beluchting na inbedrijfname grootschalig gerenoveerd of gemodificeerd? PB_LOC_BS1_unit1_isBeluchtingGemodificeerd	Aggregeren tot locatieniveau. Als een van de beluchtingsstappen gemodificeerd is dan is de waarde voor de locatie gelijk aan TRUE.	ja	14xTRUE 26xFALSE
Waterkwaliteit effluent beluchting	De waterkwaliteit van het effluent van de beluchting wordt in de meeste gevallen niet bepaald (11 van de 43 is bekend). In deze analyse nog geen focus vanwege het beperkt aantal data, eventueel te gebruiken als verklarende informatie bij gevonden verband.	-	
PARAMETERS FILTRATIE			
Heeft de locatie een ruwwaterkelder? PB_LOC_FT_isRuwWaterKelder	Slechts 2 van de 43 locaties hebben een ruwwater kelder.	ja	
PB_LOC_FT_RuwWaterKelder_IsBelucht	Niet gebruikt in analyse, te weinig data.	-	
Aantal filtratiestappen in serie. PB_LOC_FT_StapinSerie	In de meeste gevallen geldig op locatieniveau. In sommige gevallen vindt parallel filtratie over 1 stap en over 2 stappen plaats. In dat geval is het aantal stappen in serie op locatieniveau gelijk gesteld aan 2.	ja	allen
Aantal parallelle filterstraten, bij 1 of 2 filterstappen in serie, in geval dat het aantal voorfilters en nafiltsers wel of niet aan elkaar gelijk zijn. PB_LOC_FT_StapParallel	Aggregeren tot locatie niveau: Evident als aantal VF gelijk aan aantal NF (eventueel verdeeld over 2 aparte units). Wanneer aantal VF en NF niet gelijk aan elkaar dan aantal parallelle straten bepalen door te middelen (5+4 wordt 5; 6+4 wordt 5; 18+10 wordt 14; 8+4 wordt 6; 3+6 wordt 5; 3+2 wordt 3)	ja	allen
PB_LOC_FTx_unitY_FiltStraat_AantalParallel_Als_1_FTinSerie			
PB_LOC_FTx_unitY_FiltStraat_AantalParallel_Als_2_FTinSerie	In deze analyse nog geen focus op verschil tussen locaties met een verschillend aantal voor- en nafiltsers, eventueel te gebruiken als verklarende informatie bij gevonden verband.	-	
PB_LOC_FTx_unitY_FiltStraat_AantalParallel_Als_3_FTinSerie			
Is er een tussenkelder? PB_LOC_FT_FiltStraat_isKelderTussenFilterstappen	Geldt op locatieniveau. 12 locaties met een tussenkelder, 12 locaties zonder tussenkelder.	ja	12xTRUE 12xFALSE
Het procentuele debiet over een filterstraat in geval dat er meerdere parallelle units zijn. PB_LOC_FT1_unit1_DebietFiltratie%	Gebruikt in berekeningen.	-	allen
Het minimale, gemiddelde en maximale debiet over een filtersectie. PB_LOC_FT1_unit1_ProdDebietUur_FiltratieSectie_Max/Gem/Min	Gebruikt in berekening voor de ratio tussen gem en max debiet.	-	allen
Verhouding tussen het gemiddelde en maximale filtratiedebiet PB_LOC_FT_ProdDebietUur_FiltratieSectie_Ratio_MaxGem	Geaggregeerd op locatieniveau op basis van gewogen debiet per filtersectie	ja	allen
Is het productiedebiet gedurende een dag (nagenoeg) constant, bv door gebruik van vraag voorspellingssoftware? PB_LOC_FT_ProdDebietUur_isConstant	18 locaties hebben een nagenoeg constant debiet, en 20 locaties hebben een variabel debiet over de dag.	ja	18xTRUE 20xFALSE
Minimaal, gemiddeld en maximaal filtratiedebiet	Gebruikt voor berekening van filtratiesnelheid en	-	allen

per filter. PB_LOC_FT1_unit1_ProdDebietUur_FilterUnit_Max	contacttijd. Gemiddeld debiet gebruiken voor aanvullende correlatie analyse	ja	
Ratio tussen max en gemiddeld debiet op filterniveau: PB_LOC_FTx_unitY_ProdDebietUur_FilterUnit_Ratio_MaxGem	De waarde van deze parameter is vaak vergelijkbaar met de waarde van de parameter op sectie niveau. Geen extra correlatie analyse mee uitgevoerd. Eventueel te gebruiken als verklarende informatie bij gevonden verband.	-	
Dagproductie van een filtersectie PB_LOC_FTx_unitY_ProdDAG_FiltratieSectie_Max/Gem/Min	Geen focus. Informatie over ratio tussen maximaal, gemiddelde en minimale productie uit andere parameters halen, evenals de inschatting of de productie constant is gedurende de dag. Eventueel te gebruiken als verklarende informatie bij gevonden verband.	-	32
Filters continu in bedrijf? PB_LOC_FTisContinuInBedrijf	Meestal hetzelfde voor verschillende filterstappen in een locatie, daarmee ligt FALSE of TRUE voor hele locatie vast. Als verschillend, dan keuze voor gehele locatie gelijk aan waarde van laatste filterstap nemen.	ja	16xTRUE 29xFALSE
Worden de filters dagelijks geschakeld, en hoe vaak? PB_LOC_FT1_unit1_FTisDagelijksSchakel PB_LOC_FTisAantalDagelijkseSchakeling	De meeste filters worden dagelijks geschakeld. Niet meenemen in de analyse, niet onderscheidend. Analyse uitvoeren voor de schakelfrequentie (aantal schakelingen gelijk voor alle filters binnen een locatie)	ja ja	30xTRUE 0xFALSE 21
Worden filters geleidelijk of ineens in bedrijf genomen? PB_LOC_FTisIneensInBedrijf Worden de filters na een spoeling ineens of geleidelijk in bedrijf genomen? PB_LOC_FT_isIneensInBedrijfNaSpoelen	Altijd gelijk voor alle filters binnen een locatie. Sterke relatie tussen inbedrijfname (na stop) en inbedrijfname na spoeling. Binnen enkele locaties zijn er filters die geleidelijk en filters die ineens in bedrijf worden genomen na een spoeling. In die gevallen wordt de waarde voor de gehele locatie gelijk gesteld aan "geleidelijk". Analyse uitvoeren voor parameter "na spoeling".	ja	35xTRUE 7xFALSE
Is er sprake van een carroussel bedrijf? PB_LOC_FT_unit_FTisCarrousselBedrijf	Bij 4 locaties worden filters in carroussel bedreven, bij 39 locaties niet.	ja	4xTRUE 39xFALSE
Is de bovenwaterstand constant of oplopend? PB_LOC_BovenwaterstandFilt_isConst :	Vaak gelijk voor alle filters binnen een locatie. Indien verschillend, dan wordt de keuze voor gehele locatie gelijk stellen aan de waarde van laatste filterstap.	ja	29xTRUE 12xFALSE
Wat is de bovenwater stand (constant of max)? PB_LOC_FTx_unitY_Bovenwaterstand_Constant_Hoogte PB_LOC_FTx_unitY_Bovenwaterstand_Oplopend_MaxHoogte	De waarde wordt gebruikt in berekeningen, o.a. van de contacttijd.	-	24
Wat is de leeftijd van de verschillende filtermedia? PB_LOC_FT_isLeeftijdMedium	De leeftijd van de verschillende filterlagen en van de verschillende filters worden gemiddeld om tot een locatie geaggregeerde waarde te komen.	ja	
De ondermaat en de bovenmaat van filtermedia PB_LOC_FTx_unitY_LaagX_isOndermaat PB_LOC_FTx_unitY_LaagX_isBovenmaat	In deze analyse nog geen focus vanwege het beperkt aantal data, eventueel te gebruiken als verklarende informatie bij gevonden verband.	-	
Zijn de filters enkel of dubbellaags? PB_LOC_FTx_unitY_FT_isEnkellaags	Vaak is de waarde gelijk voor alle filters binnen een locatie. Als 1 TRUE en 1 FALSE, dan keuze voor gehele locatie gelijk stellen aan dubbellaags (FALSE) filter (daarmee wordt de vraag: heeft de locatie alleen maar enkellaag filtratie?). In andere gevallen kiezen voor het merendeel (bv Noordbergum = 2 x TRUE en 1 x FALSE).	ja	
Hebben de filters een steunlaag? PB_LOC_FTx_unitY_FT_isSteunlaag :	Vaak zijn de waarden gelijk binnen 1 locatie. Als de waarden verschillend zijn, dan is de vraag geïnterpreteerd	ja	

	als "is er een filter in de locatie die een steunlaag heeft?" Als 1 filter een steunlaag heeft wordt dat ook de waarde voor de gehele locatie.		
Is er sprake van gravitatie of druk filtratie? PB_LOC_FT _x _unitY_FT_isGravitatie	Vaak is de waarde gelijk voor alle filters binnen een locatie. Als niet, dan is de keuze voor de gehele locatie gelijk gesteld aan de laatste filterstap.	ja	
Is er sprake van droog of nat filtratie? PB_LOC_FT_isDroog	Vaak is de waarde gelijk voor alle filters binnen een locatie. Als er een mix is (1:1), dan is de vraag geïnterpreteerd als 'is één van de filters een droog filter?'. Als er een mix is maar het merendeel van de filters is nat (bv Helmond, 3:1), dan als wordt de waarde van de locatie gelijk gesteld aan nat filtratie.	ja	
Wordt het bed belucht tijdens filtratie? PB_LOC_FT_BedBelucht Het luchtdebiet van deze beluchting: PB_LOC_FT1_unit1_FT_BedBeluchting_debietTOT	Vaak is de waarde van verschillende filters gelijk voor een locatie. Als er sprake is van een mix (1:1), dan is de vraag geïnterpreteerd als 'wordt één van de filters belucht?'. Als er sprake is van een mix maar het merendeel wordt niet belucht (bv Helmond, 3:1), dan als is de waarde van de locatie gelijk aan 'niet belucht'. In deze analyse nog geen focus gelegd op het luchtdebiet, eventueel te gebruiken als verklarende informatie bij gevonden verband.	nee	
Filtermedium: grind/zand, marmer, antraciet of anders.	Veelal grind (42 van de 43). Daarom niet meenemen in de analyse. Als een ander medium dan grind, dan is dat vaak antraciet, en dan in combinatie met een dubbellaags filter. Eventueel te gebruiken als verklarende informatie bij gevonden verband.	-	
Verdere informatie over het filtermedium: - Slijtage - Aangroei - Eventuele reiniging - Conditie van het filter	In deze analyse nog geen focus gelegd op het luchtdebiet, eventueel te gebruiken als verklarende informatie bij gevonden verband.	-	
Waterkwaliteit effluent filtratie	De waterkwaliteit van het effluent van de filters is voor verschillende productielocaties bekend. In deze analyse nog geen focus, eventueel te gebruiken als verklarende informatie bij gevonden verband. Opletten voor beperkt aantal data bij inlopen filters (bv hoge KG22)	-	
Aantal spoelgoten PB_LOC_FT_AantalSpoelgoot	Voor de meeste locaties is het aantal spoelgoten voor alle filters gelijk.	ja	
Positie spoelgoot (zijkant/midden en korte/lange kant) - PB_LOC_FT_SpoelgootZijkant - PB_LOC_FT_SpoelgootKorteKant	Voor de meeste filters binnen een locatie gelijk. Indien verschil, dan: - Waarde locatie gelijk aan 'zijkant'. - Merendeel voor korte/lange kant	ja	
Afstand van filterbed tot aan spoelgoot PB_LOC_FT_AfstandSpoelgoot_Filterbed	Aggregeren door gemiddelde te nemen indien de afstand per filter verschillend is.	ja	
Aantal spoeldoppen per m2 PB_LOC_FT_AantalSpoeldoppen	Aggregeren door gemiddelde te nemen indien deze waarde per filter verschillend is.	ja	

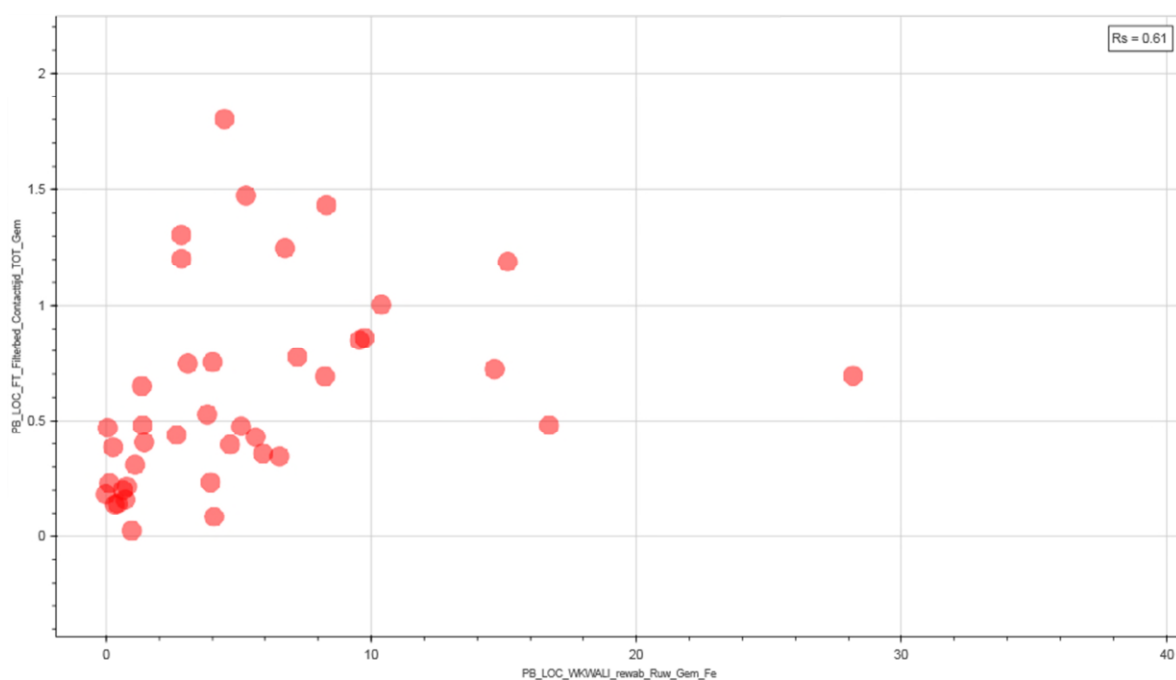
Luchtinbreng bij spoelen op 1 punt onder filter of verdeeld? PB_LOC_FT1_isLuchtinbreng_1Punt	Is voor elke locatie gelijk voor alle filters.	ja	
Spoellucht behandeld? PB_LOC_FT1_Spoellucht_isBehandeld	Is voor elke locatie gelijk voor alle filters. Aanvullende gegevens zoals type behandeling: In deze analyse nog geen focus gelegd op het luchtdebiet, eventueel te gebruiken als verklarende informatie bij gevonden	Ja -	
Spoelcriterium (tijd, volume, waterkwaliteit, drukval)	In deze analyse nog geen focus gelegd op het luchtdebiet, eventueel te gebruiken als verklarende informatie bij gevonden. Wel analyses uitvoeren met de looptijd.	-	
Looptijd filter PB_LOC_FT_x_unitY_Filt_Looptijd	Gebruikt bij berekening spoelwater volume voor locatie.	-	
Genormaliseerd spoelwatervolume en spoellucht volume (genormaliseerd naar filteroppervlak) PB_LOC_FT_x_unitY_TotNormVolSpoelwater PB_LOC_FT_x_unitY_TotNormVolSpoelLucht	Gebruikt bij berekening spoelwater volume voor locatie. Berekend op basis van onderliggende informatie per spoelwaterfase (som debiet spoeling x duur spoelfase / oppervlak filter). Uitgedrukt in Ook analyse op locatie geaggregeerd niveau (debiet gewogen gemiddelde over parallelle units, waarden in serie opgeteld).	ja ja	
Wordt de spoeling handmatig of middels spoelautomaat uitgevoerd?	Niet geanalyseerd, nauwelijks onderscheidend	-	
Maximale spoelsnelheid (water en lucht) PB_LOC_FT_MaxSpoelsnelheidWater PB_LOC_FT_MaxSpoelsnelheidLucht	Geaggregeerd tot locatie niveau (debiet gewogen gemiddeld)	ja	
Is er sprake van eerste filtraat afvoer? PB_LOC_FT_isEersteFiltraatAfvoer	Geaggregeerd tot locatie niveau. Voor de meeste locaties voor alle filterstappen gelijk binnen een locatie. Indien verschillend, dan waarde locatie gelijk stellen aan de waarde van de laatste filterstap.	ja	
Aanvullende informatie over spoelprogramma's: - Geoptimaliseerd? Bijvoorbeeld met uitspoelcurve? - Wanneer? - Aanwijzingen voor niet optimaal spoelprogramma?	In deze analyse nog geen focus gelegd op deze parameters, eventueel te gebruiken als verklarende informatie bij gevonden.	-	
Wordt het filtermedium regelmatig schoongemaakt? PB_LOC_FT_isPeriodiekeSchoonmaakFT	Aggregeren op locatie niveau. Meestal gelijke waarde voor verschillende filters binnen een locatie. Indien verschil, dan waarde locatie gelijk stellen aan TRUE (schoonmaak), of aan het merendeel van de waarden.	ja	
Aanvullende informatie over schoonmaak van filters - Schoonmaak op basis van welk criterium? - Hoe vaak schoonmaken? - Wat houdt schoonmaak in?	In deze analyse nog geen focus gelegd op deze parameters, eventueel te gebruiken als verklarende informatie bij gevonden.	-	
Filtratiesnelheid (gemiddeld en maximaal)	Berekend op basis van filtratie debieten en filteroppervlak.	ja	

	Geaggregeerd op locatieniveau op basis van debiet gewogen gemiddelde.		

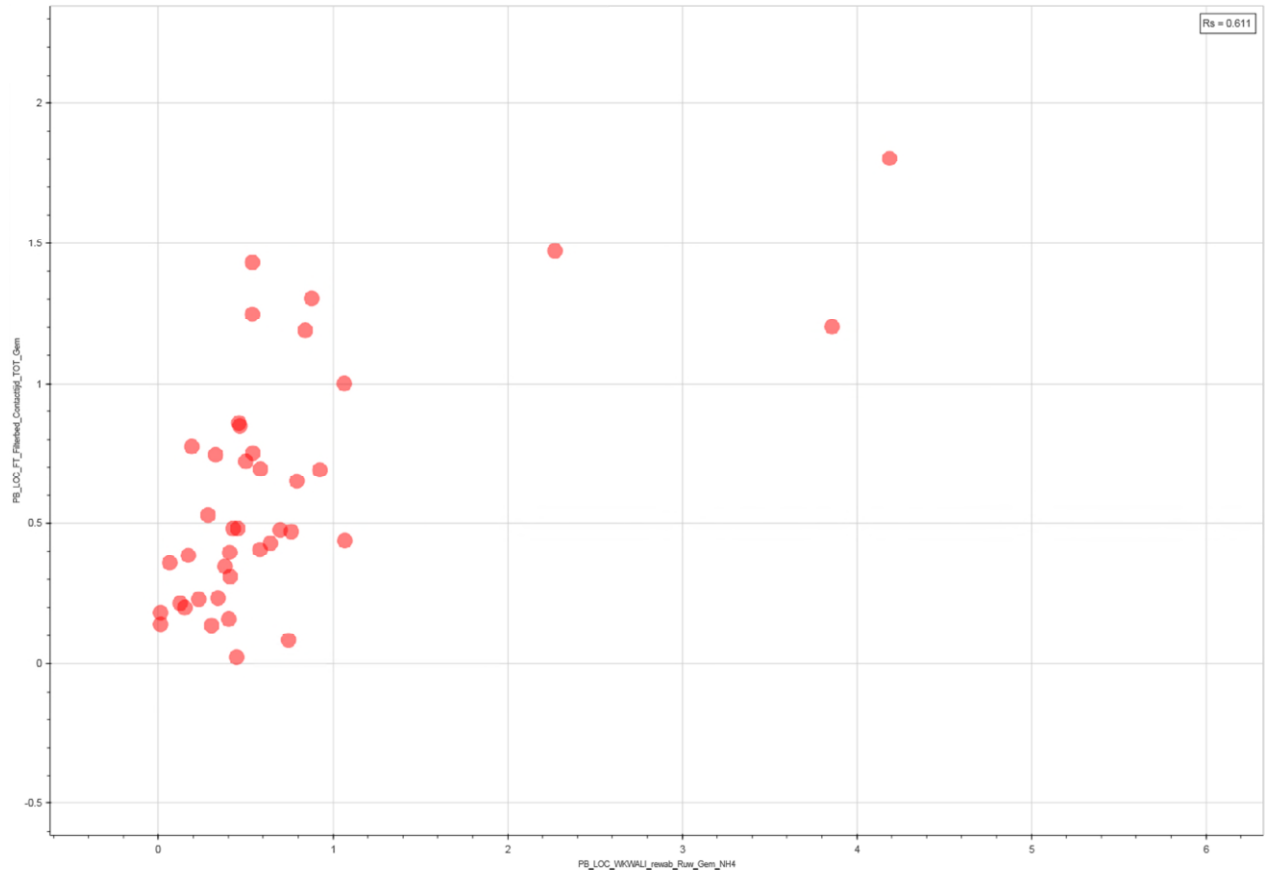
Bijlage II Additionele relaties

Relaties ruwwater en bedrijfsvoering

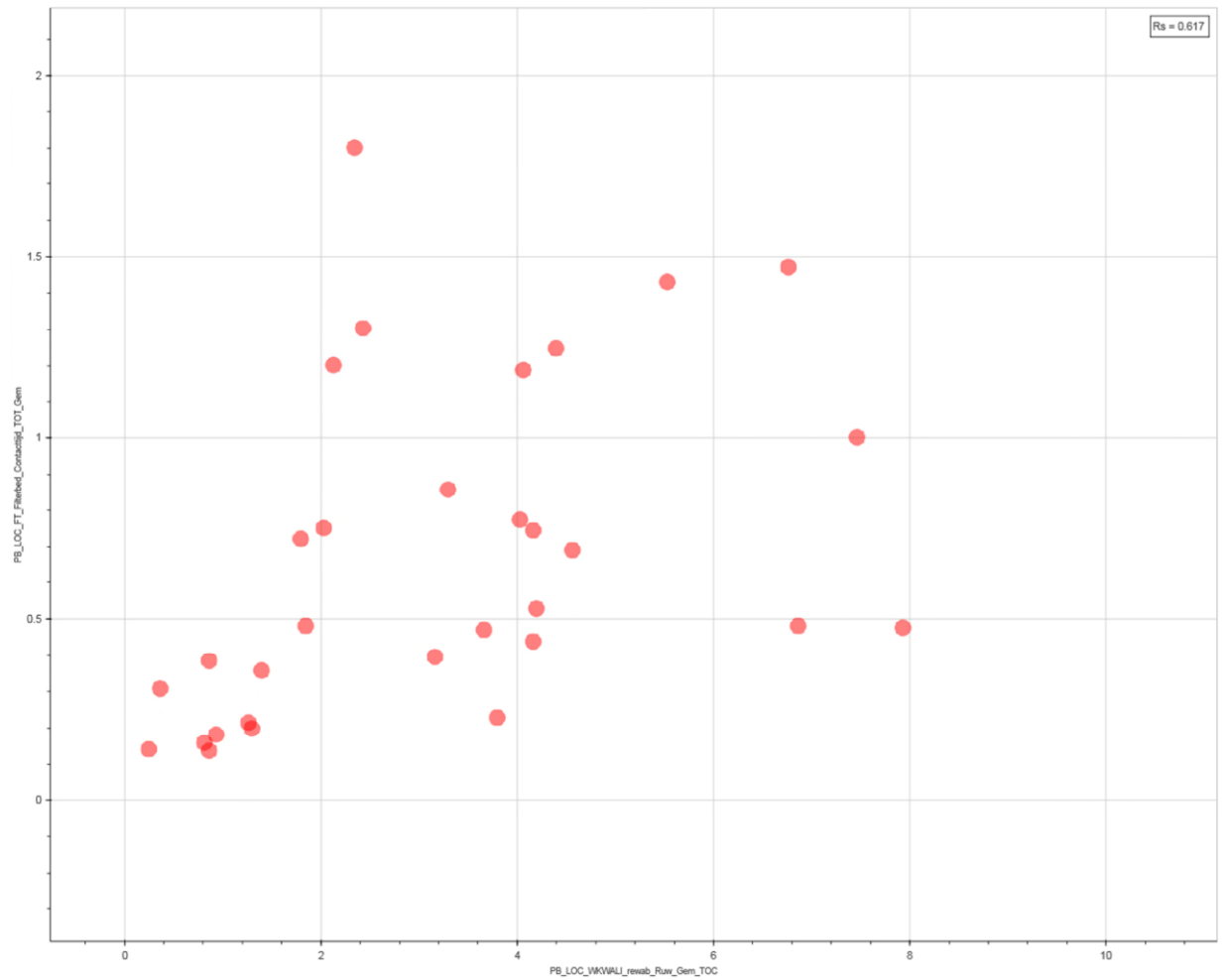
Filterbed contacttijd (uur) versus ijzer ruwwater (Spearman coëfficiënt 0,61), zie figuur. De filterbed contacttijd correleert tevens met de verwijdering van ijzer (Spearman coëfficiënt is 0,51 voor de conversie van ijzer).



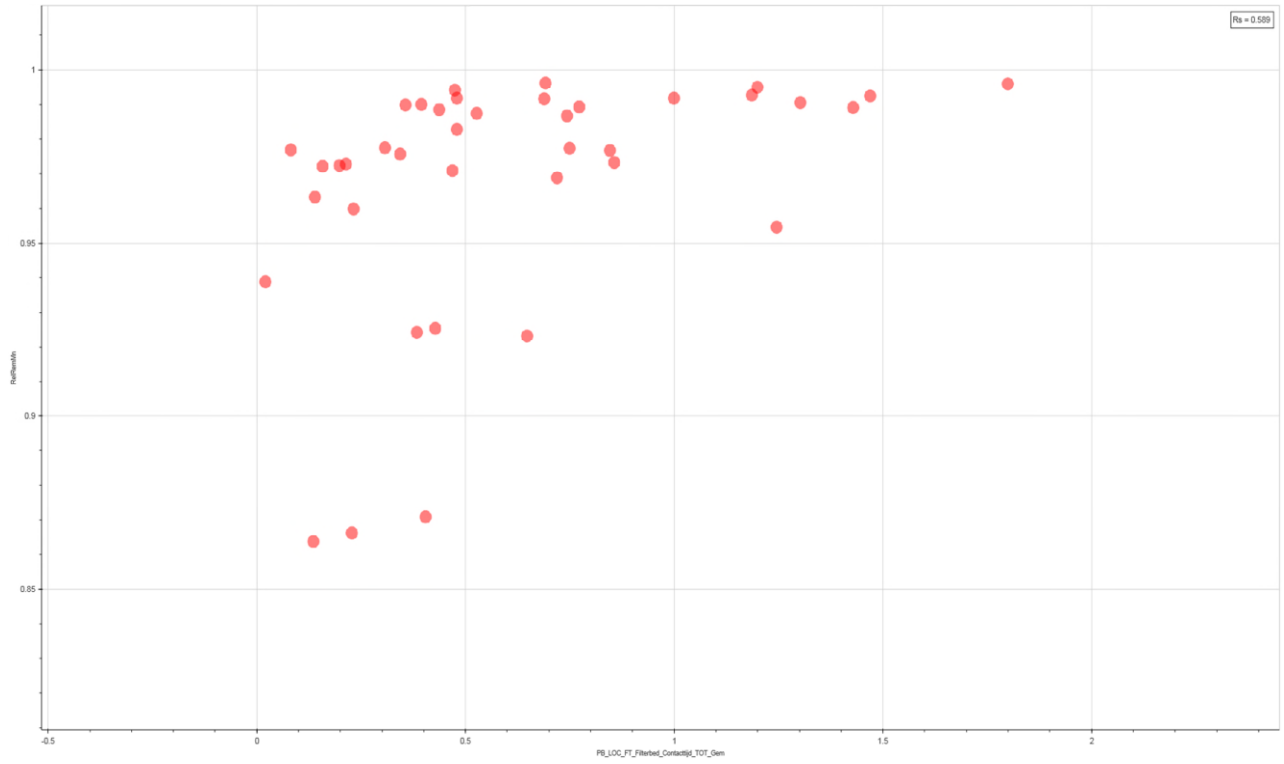
Filterbed contacttijd (uur) versus ammonium ruwwater (mg/L)



Filterbed contacttijd (uur) versus TOC ruwwater (mg/L)

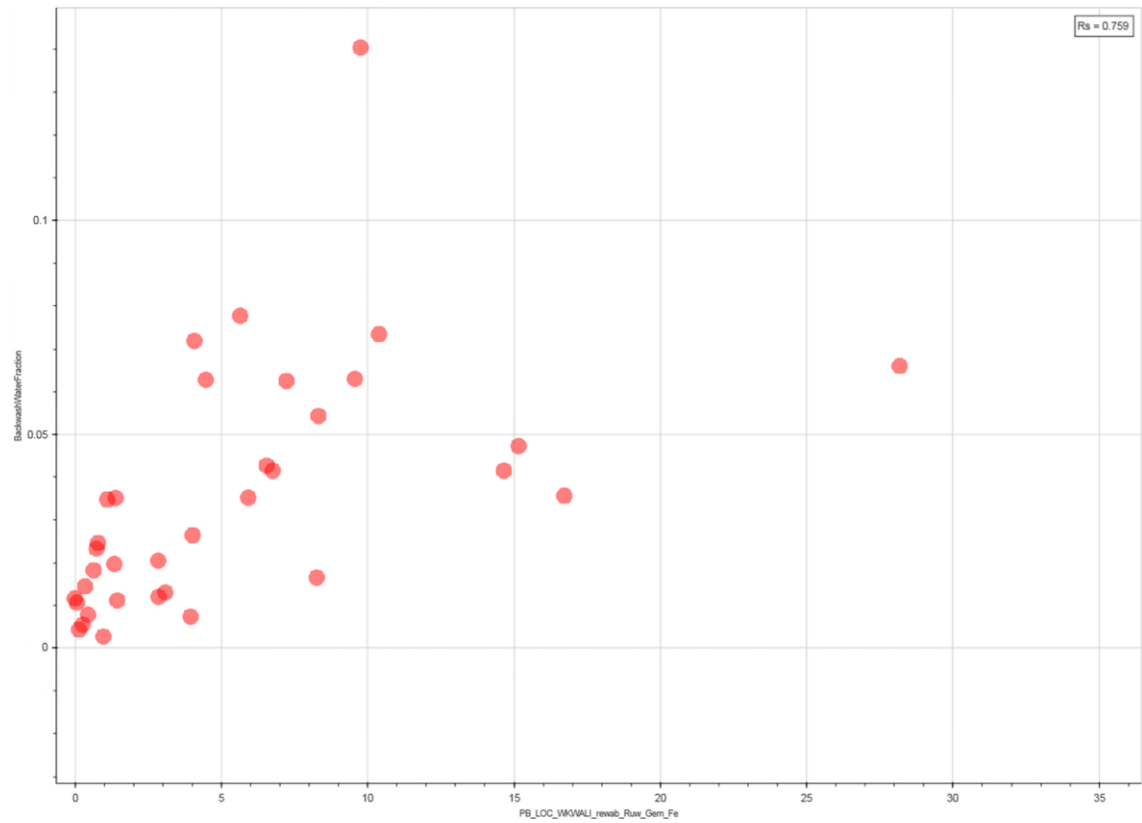


Conversie mangaan (%) versus filterbed contacttijd (uur)

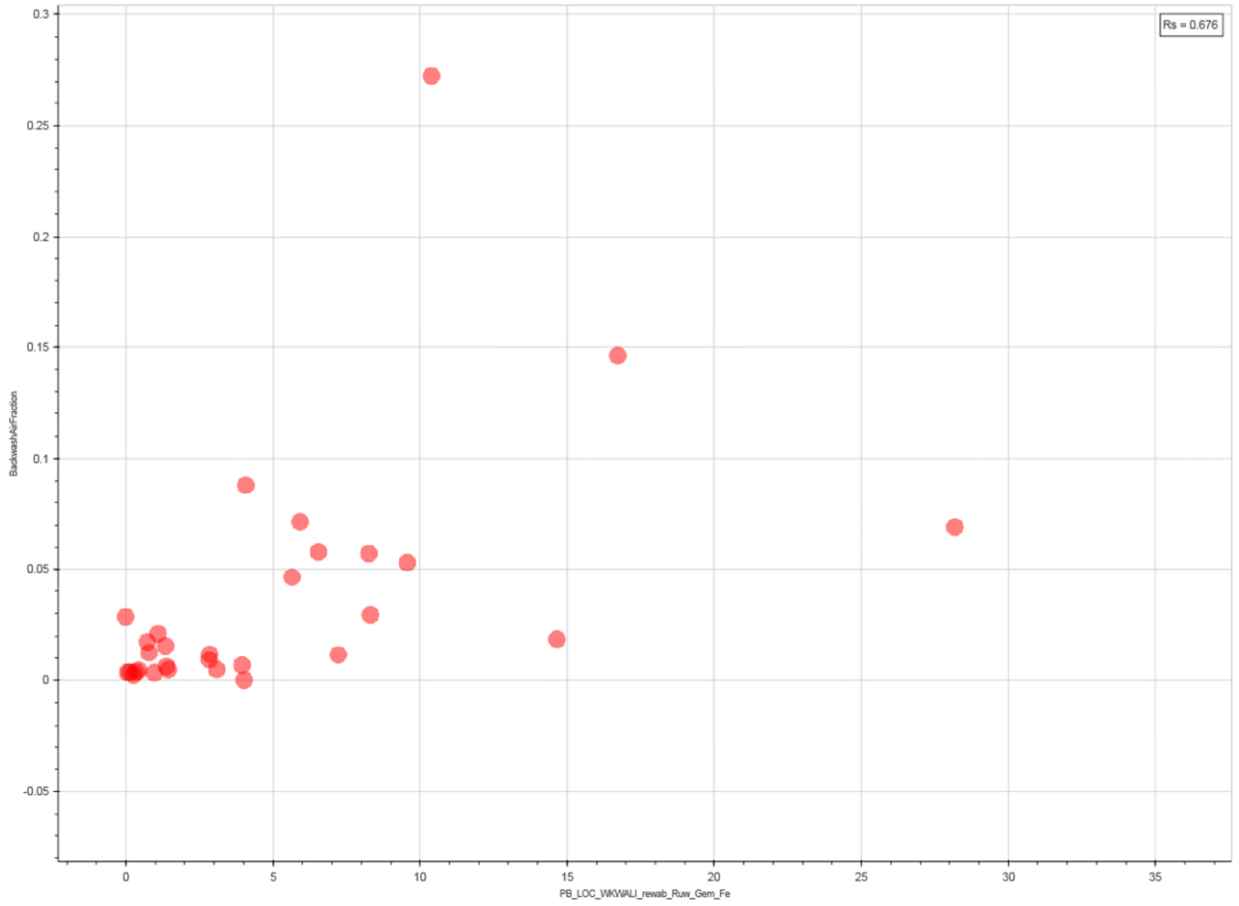


Spiegelwaterpercentage versus ruwwater ijzergehalte (mg/L). Spearman coëfficiënt 0,76.

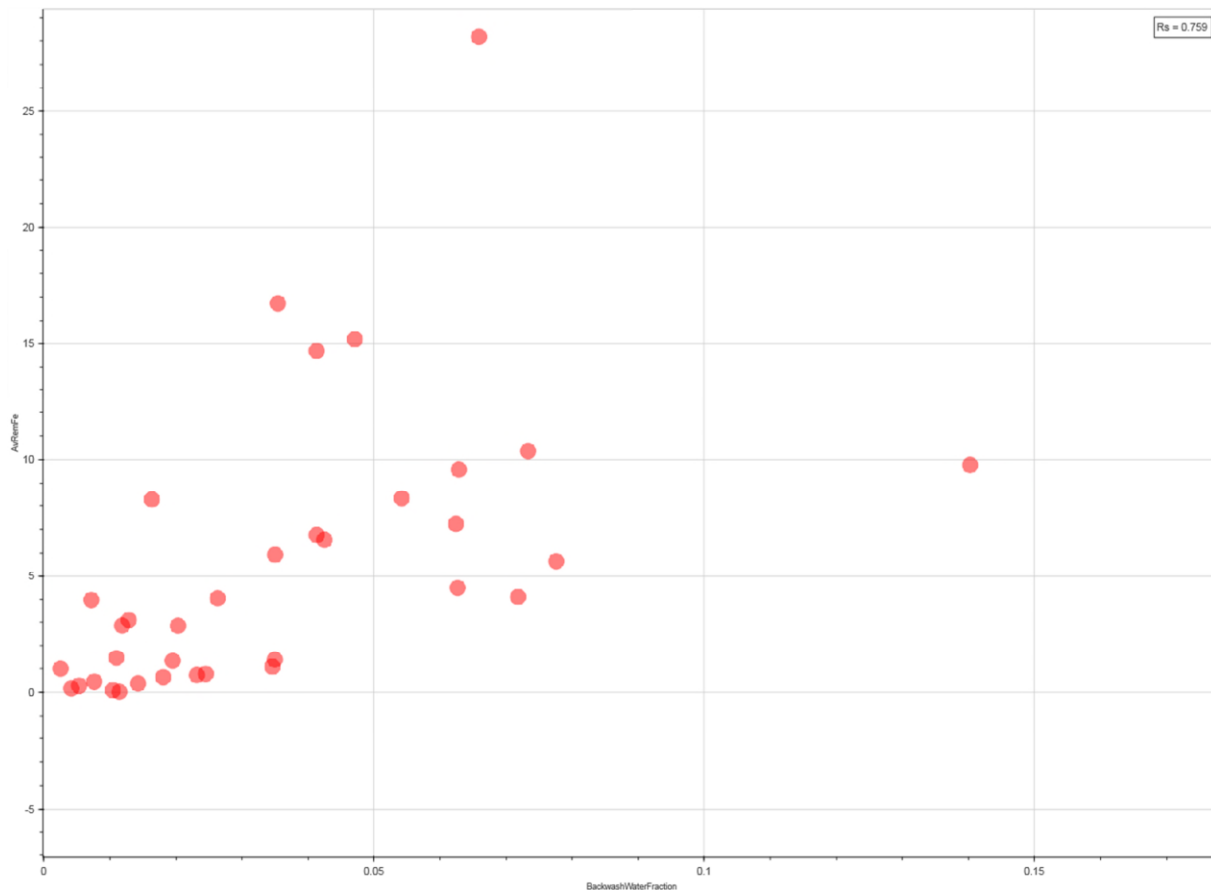
Vergelijkbaar verband voor ruwwater mangaan (0,57).



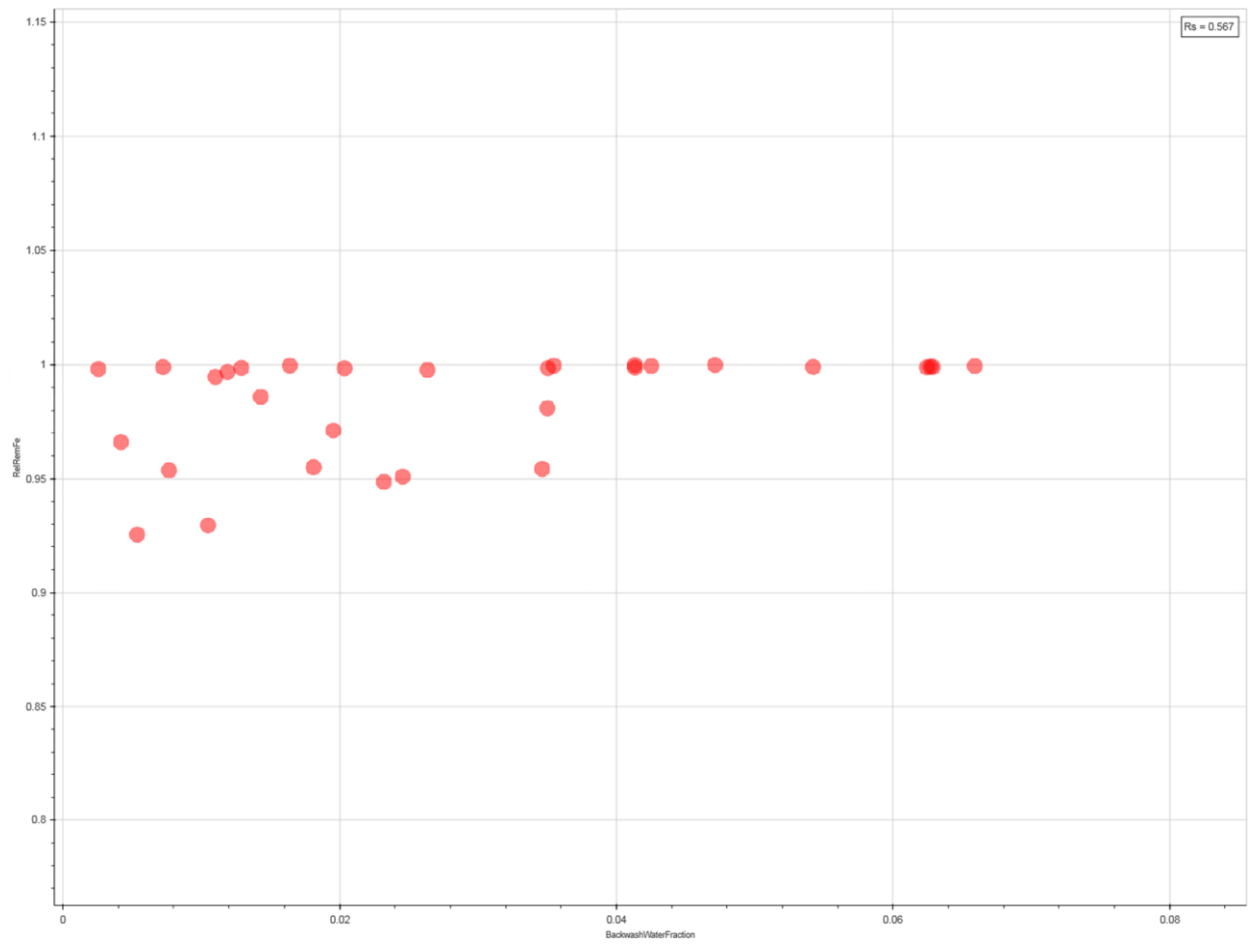
Spoelluchtpercentage versus ruwwater ijzergehalte (mg/L). Spearman coëfficiënt 0.68.
 Vergelijkbaar verband voor ruwwater mangaan (0,6).



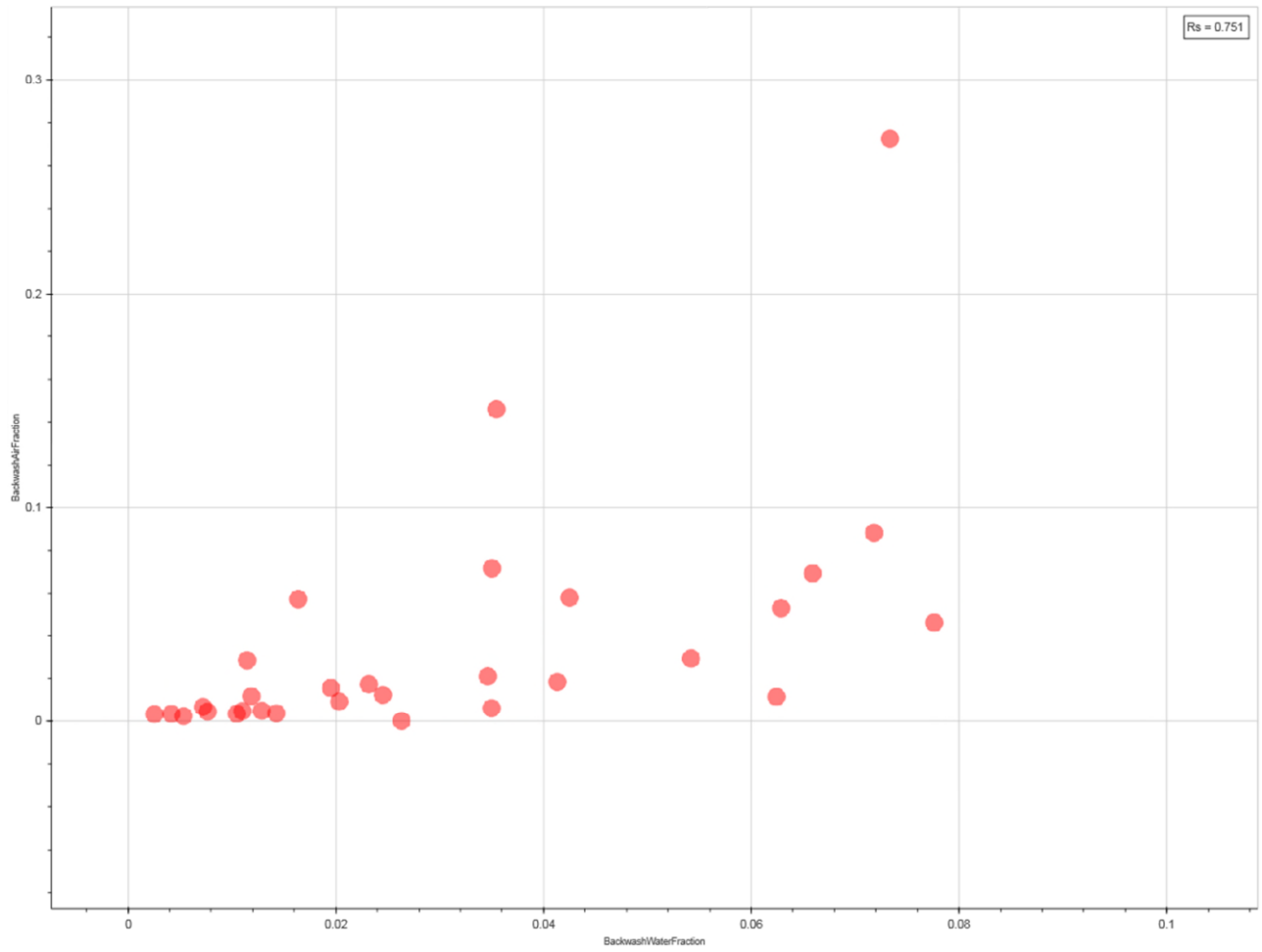
Spoelwaterpercentage versus (absolute) verwijdering ijzer (mg/L). Spearman coëfficiënt 0,76. Een vergelijkbaar verband is gevonden voor het spoelluchtpercentage en de verwijdering van ijzer (0,68). Vergelijkbare verbanden zijn gevonden tussen de spoelwater- en spoelluchtpercentages en de absolute verwijdering van mangaan (resp. 0,56 en 0,6), en met de absolute verlaging van TOC (resp. 0,6 en 0,4).



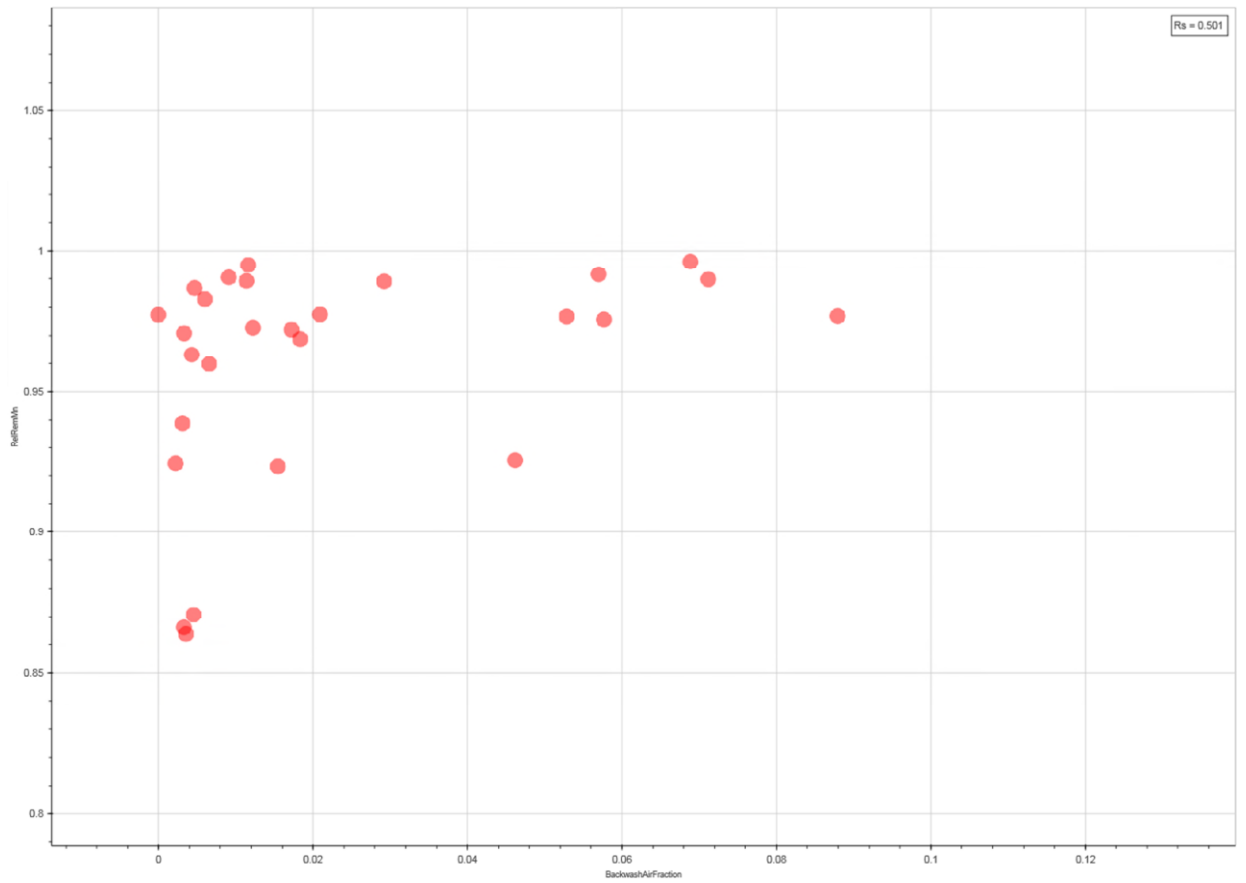
Spoelwaterpercentage versus conversie ijzer (%). Spearman coëfficiënt 0,57. Vergelijkbaar met spoellucht en ijzer (0,5).



Spoelwaterpercentage versus spoelluchtpercentage.



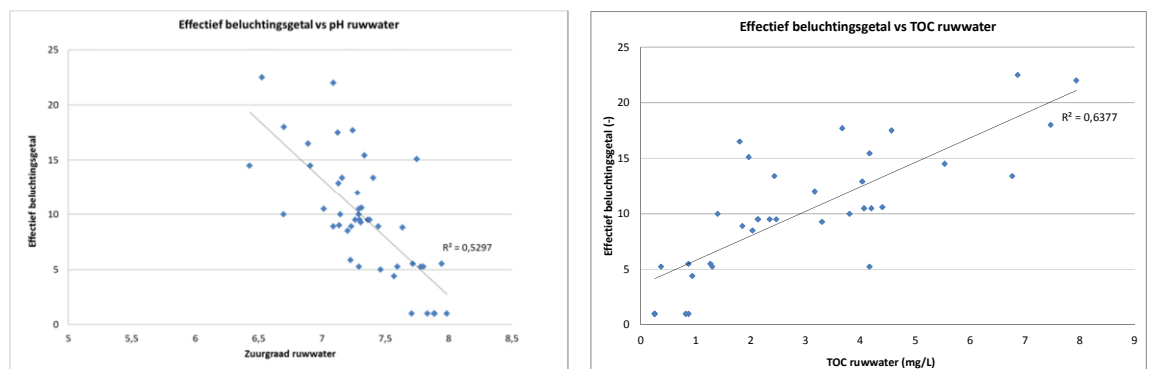
Spelwaterpercentage versus conversie mangaan (%). Spearman coëfficiënt 0,5.
 Vergelijkbaar met spoellucht en mangaan (0,5).



Relaties chemisch-fysische parameters ruwwater en conversie

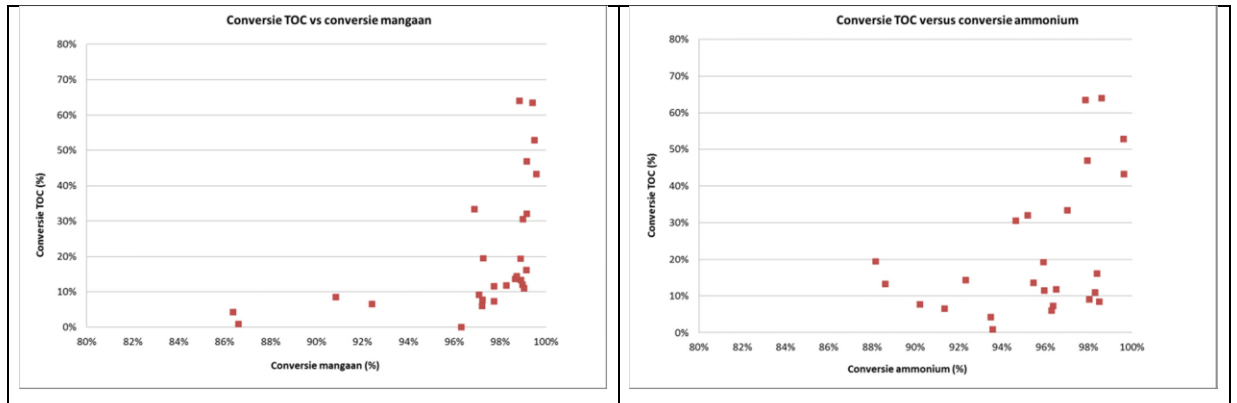
Bij het onderzoeken van de relaties met nagroei indicatoren zijn verschillende interessante verbanden gevonden tussen ruwwaterkwaliteit en de verandering van fysisch-chemische parameters. De relaties zijn hier benoemd:

- Er zijn verschillende verbanden gevonden tussen het methaan gehalte, het TOC gehalte en ammonium, o.a. in ruwwater, maar ook in het reinwater en het gedistribueerde water, voor zowel de gemiddelde als de minimum en maximum waarden.
- Er is een sterke relatie gevonden tussen de verandering van pH (procentueel) en de pH van ruwwater ($R^2 = 0,77$).
- Er is een sterke relatie gevonden tussen de verwijdering van ijzer en het ijzer gehalte in ruwwater.
- De relatie tussen de ruwwater kwaliteit en het effectief beluchtingsgetal is het sterkst voor het TOC gehalte. De relaties tussen het effectief beluchtingsgetal en de zuurgraad of methaan in ruwwater zijn ook sterk (zie Figuur 31). De relatie met de pH is ook besproken in paragraaf 3.6.1. Uit de relatie met methaan valt op dat er verschillende locaties zijn die een hoog effectief beluchtingsgetal hebben ondanks een beperkt methaan gehalte. Er is tevens een relatie gevonden tussen het effectief beluchtingsgetal en de hoeveelheid ijzer in ruwwater en de verwijdering van ijzer.



FIGUUR 31. ZUURGRAAD EN TOC RUWWATER VERSUS EFFECTIEF BELUCHTINGSGETAL

- Er zijn verschillende interessante verbanden gevonden tussen het ruwwatergehalte of de verwijdering van mangaan of ijzer enerzijds en de verlaging van TOC of ammonium anderzijds. Samengevat: een lage verlaging van TOC of ammonium gaat vaak samen met een minder vergaande verwijdering van ijzer of mangaan. Locaties met nanofiltratie of actieve kool filtratie hebben de hoogste TOC verlaging, een enkele locatie met ontharding toont ook een hoge TOC verlaging. Een voorbeeld is weergegeven in Figuur 32:



FIGUUR 32. CONVERSIE TOC VERSSUS CONVERSIE MANGAAN (LINKS) EN AMMONIUM (RECHTS)

Bijlage III Uitgangspunten data bewerking

In deze bijlage staan verschillende uitgangspunten opgesomd die gehanteerd zijn bij de kwaliteitscontrole en de verwerking van de aangeleverde formats.

Waterkwaliteit

- Indien een getal kleiner is dan de rapportagegrens is gerekend met de helft van de waarde van de rapportagegrens.
- Indien KG22 of *Aeromonas* gerapporteerd is als > 1000, dan is een waarde van 1000 gehanteerd.
- In sommige gevallen is het Koloniegetal bij 25 °C beschikbaar, hier zijn geen analyses mee uitgevoerd.
- In de meeste gevallen zijn data van 2011 – 2013 beschikbaar, in sommige gevallen alleen voor een deel van die periode.
- DOC is minder vaak bekend dan TOC, daarom is TOC in de analyse betrokken. In het geval dat een TOC waarde ontbreekt en de DOC waarde bekend is, dan is TOC gelijk aan DOC gesteld.
- Indien waterkwaliteitsgegevens van het gedistribueerde water ontbreken, is gebruik gemaakt van de Rewab database.
- In een enkel geval is een ontbrekende waterkwaliteit ingevuld op basis van de database voor eerder BTO onderzoek uit 2011.
- Voor Fikkersdries II is de ruwwater kwaliteit niet beschikbaar in de Rewab query, ook is geen kwaliteit bekend van de winvelden van deze locatie (Hemmen en Zetten). In dit geval is uitgegaan van de gemiddelde kwaliteit zoals gegeven in het Vitens document "Gelderland – kwaliteit en zuivering 2011".
- Er zijn geen analyses uitgevoerd met het KG22, *Aeromonas* en de troebelingsgraad van het ruwwater.

Bedrijfsvoering

- In sommige gevallen zijn maximale of ontwerp jaardebieten niet ingevuld. In die gevallen is de maximale of ontwerp capaciteit bepaald op basis van het uurdebiet x 22 uur/dag x 365. In sommige gevallen ontbreekt de ontwerpcapaciteit, deze is dan gelijk gesteld aan de maximale capaciteit, en vice versa. Wanneer de gemiddelde jaarcapaciteit niet gegeven is dan is hiervoor een aanname gemaakt op basis van het gemiddelde uur- of dagproductie. Deze eigenschappen komen vooral in de correlatiestudie terug in de benuttingsgraad.
- Het waterdebiet per beluchtingsstap is afgestemd op, en in sommige gevallen aangepast aan de hand van het waterdebiet in de filtratiesectie. Daarbij is gebruik gemaakt van de PFD.
- Indien het format melding maakt van gebruik van OPIR of andere prognosebesturing is aangenomen dat de debieten over de dag constant zijn.

Ontwerp

- Fikkersdries I en Fikkersdries II zijn als twee locaties ingevuld. De reinwaters worden gemengd en daarna gedistribueerd. Daarom zijn deze twee locaties verwerkt tot één locatie "Fikkersdries I + 2".
- Lekkerkerk Tiendweg en Lekkerkerk Schuwacht zijn als twee locaties ingevuld. De reinwaters worden gemengd en daarna gedistribueerd. Daarom zijn deze twee locaties verwerkt tot één locatie "Lekkerkerk Totaal".
- Helmond heeft twee ruwwaterstrengen en twee daarop aangepaste zuiveringen. De ruwwater zijn debietsproportioneel gemiddeld.
- Disturbances: NF, UV, AKF, ONTH en bijmengen ander reinwater zijn beschouwd als disturbance. De dosering van een chemische stof (FeCl_3 , KMnO_4 , beetje kalkmelk of natronloog tbv pH verandering), spoelwaterhergebruik, ondergrondse ontijzering, marmer filtratie zijn niet als disturbance beschouwd.
- Indien de productielocatie in de periode 2011 – 2013 een mutatie heeft ondergaan is de waarde van de betreffende parameters afgestemd op de situatie die het grootste deel van die periode van toepassing was.