

VERGELIJKENDE ANALYSE MILIEUINDICATIEWAARDEN MACROFYTEN



RAPPORT

2016

W03

VERGELIJKENDE ANALYSE
MILIEUINDICATIEWAARDEN MACROFYTEN

RAPPORT

2016
W03



COLOFON

UITGAVE Stichting Toegepast Onderzoek Waterbeheer
Postbus 2180
3800 CD Amersfoort

AUTEUR(S) **Nico Jaarsma**
Ecologie en Fotografie

DRUK Kruyt Grafisch Adviesbureau
STOWA STOWA 2016-W03

COPYRIGHT Teksten en figuren uit dit rapport mogen alleen worden overgenomen met bronvermelding.

DISCLAIMER Deze uitgave is met de grootst mogelijke zorg samengesteld. Niettemin aanvaarden de auteurs en de uitgever geen enkele aansprakelijkheid voor mogelijke onjuistheden of eventuele gevolgen door toepassing van de inhoud van dit rapport.

SAMENVATTING

Planten en dieren stellen bepaalde eisen aan hun milieu, het voorkomen van een bepaalde soort kan daarom iets zeggen over de milieu-omstandigheden. Dit is milieu-indicatie. In de wetenschap is van oudsher veel gewerkt aan milieu-indicatie, bekend zijn vooral de Ellenberg-getallen voor o.a. zuurgraad en productiviteit van de bodem op basis van vegetatie. Er zijn verschillende methoden om milieu-indicatiewaarde af te leiden en er zijn ook al bestaande sets van indicatiewaarden, o.a. voor aquatische vegetatie.

Deze studie gaat over de vraag of en hoe vegetatiegegevens kunnen worden gebruikt om een (kwantitatieve) uitspraak te kunnen doen over het milieu. Bijvoorbeeld het voorspellen van de zuurgraad of de fosfaatrijkdom op basis van een lijstje plantensoorten. Belangrijke vragen daarbij zijn:

- Kan dit? Zo ja, hoe goed is de voorspelling?
- Welke rekenmethoden zijn dan het best bruikbaar?
- Moeten soorten- en milieudata nog worden getransformeerd?
- Op welk niveau werken indicatiewaarden het best? Landelijk, per watertype, per regio?

Het project heeft ten doel om inzichtelijk te maken welke keuzes gemaakt moeten worden om te komen tot een geschikte set aan milieu-indicatiewaarden.

In dit rapport zijn twee verschillende methoden om milieu-indicatiewaarden te bepalen onderling vergeleken: weighted averaging (berekening van optima en toleranties volgens methode ter Braak) en gewogen gemiddelden (volgens methode Bloemendaal & Roelofs). Hoewel de termen “weighted averaging” en “gewogen middelen” feitelijk hetzelfde betekenen, is de berekeningswijze voor beide methoden duidelijk verschillend. Ze leiden ook tot duidelijk andere resultaten, zoals in dit rapport wordt geïllustreerd. Voor de afleiding van de indicatiewaarden zijn gegevens van macrofyten en milieucondities uit de Limnodata gebruikt. De resultaten zijn vergeleken met bestaande indicatiewaarden voor macrofyten van AqMaD en van De Lyon & Roelofs en met indicatiewaarden voor macrofyten die zijn afgeleid voor Waternet en Waterschap Rivierenland.

De indicatiewaarden zijn berekend voor de gehele dataset (macrofyten van stilstaande wateren) en voor deelsets daarvan. Bij de deelsets is onderscheid gemaakt tussen: landschap-ecologische regio's volgens OBN (laagveen, zeeklei, rivierengebied, hogere zandgronden) en hoofdwatertypen (sloten, kanalen, meren en plassen, diep en ondiep). Ook is nog onderscheid gemaakt in beheergebieden van (enkele grotere) waterbeheerders: Hoogheemraadschap Hollands Noorderkwartier (HHNK), Waterschap Hollandse Delta (WHD), Waterschap Rivierenland (WSRL) en Waterschap Scheldestromen (WSS).

De uitkomsten laten zien dat de indicatiewaarden van soorten in sterke mate worden beïnvloed door bovengenoemde keuzes. Zowel de keuze van de regio als van het watertype heeft een duidelijke invloed op de berekende indicatiewaarden. Ook is de mate waarin de meetgegevens al dan niet evenwichtig zijn verdeeld over de milieugradiënt hierop van (grote) invloed. Dit geldt voor de absolute indicatiewaarden. In veel gevallen was de ordening van indicatiewaarden (relatief ten opzichte van elkaar) in de verschillende datasets wel

vergelijkbaar. Hierdoor is een “relatieve” toepassing van de indicatiewaarden, zoals in AqMaD, met de bestaande data en methoden goed mogelijk, ondanks de verschillen tussen regio's en watertypen.

Aan een “absolute” toepassing (zoals het voorspellen van de waarde van een milieuvariabele vanuit de soorten) zitten verschillende haken en ogen. Hiertoe is een alternatieve methode bedacht en onderzocht die uitgaat van de kans op voorkomen van soorten over een milieugradiënt. De methode kan omgekeerd worden gebruikt voor milieu-indicatie.

Aanbevolen wordt om zowel alternatieve methoden nader te bekijken als om datasets te verbeteren en - voorafgaand aan het bepalen van indicatiewaarden - in te delen in enkele grotere clusters op basis van regio en watertype.

DE STOWA IN HET KORT

STOWA is het kenniscentrum van de regionale waterbeheerders (veelal de waterschappen) in Nederland. STOWA ontwikkelt, vergaart, verspreidt en implementeert toegepaste kennis die de waterbeheerders nodig hebben om de opgaven waar zij in hun werk voor staan, goed uit te voeren. Deze kennis kan liggen op toegepast technisch, natuurwetenschappelijk, bestuurlijk-juridisch of sociaalwetenschappelijk gebied.

STOWA werkt in hoge mate vraaggestuurd. We inventariseren nauwgezet welke kennisvragen waterschappen hebben en zetten die vragen uit bij de juiste kennisleveranciers. Het initiatief daarvoor ligt veelal bij de kennisvragende waterbeheerders, maar soms ook bij kennisinstellingen en het bedrijfsleven. Dit tweerichtingsverkeer stimuleert vernieuwing en innovatie.

Vraaggestuurd werken betekent ook dat we zelf voortdurend op zoek zijn naar de 'kennisvragen van morgen' – de vragen die we graag op de agenda zetten nog voordat iemand ze gesteld heeft – om optimaal voorbereid te zijn op de toekomst.

STOWA ontzorgt de waterbeheerders. Wij nemen de aanbesteding en begeleiding van de gezamenlijke kennisprojecten op ons. Wij zorgen ervoor dat waterbeheerders verbonden blijven met deze projecten en er ook 'eigenaar' van zijn. Dit om te waarborgen dat de juiste kennisvragen worden beantwoord. De projecten worden begeleid door commissies waar regionale waterbeheerders zelf deel van uitmaken. De grote onderzoekslijnen worden per werkveld uitgezet en verantwoord door speciale programmacommissies. Ook hierin hebben de regionale waterbeheerders zitting.

STOWA verbindt niet alleen kennisvragers en kennisleveranciers, maar ook de regionale waterbeheerders onderling. Door de samenwerking van de waterbeheerders binnen STOWA zijn zij samen verantwoordelijk voor de programmering, zetten zij gezamenlijk de koers uit, worden meerdere waterschappen bij één en het zelfde onderzoek betrokken en komen de resultaten sneller ten goede van alle waterschappen.

De grondbeginselen van STOWA zijn verwoord in onze missie:

Het samen met regionale waterbeheerders definiëren van hun kennisbehoeften op het gebied van het waterbeheer en het voor én met deze beheerders (laten) ontwikkelen, bijeenbrengen, beschikbaar maken, delen, verankeren en implementeren van de benodigde kennis.

VERGELIJKENDE ANALYSE MILIEUINDICATIEWAARDEN MACROFYTEN

INHOUD

	TEN GELEIDE SAMENVATTING DE STOWA IN HET KORT	
1	INLEIDING	1
1.1	Aanleiding	1
1.2	Habitat-eisen van macrofyten	1
1.3	Centrale vraag en doelstelling	2
1.4	Datasets en uitgevoerde analyses	2
1.5	Leeswijzer	3
2	VOORBEWERKING EN ANALYSE	4
2.1	Verzamelen en voorbereiden data	4
2.1.1	Bestaande indicatiewaarden	4
2.1.2	Afleiden indicatiewaarden uit de Limnodata	5
2.2	Sturende factoren vegetatie en selectie parameters voor analyse	6
2.2.1	Sturende factoren	6
2.2.2	Selectie parameters voor nadere analyse	9
2.3	Overzicht datasets en uitgevoerde analyses	9
2.3.1	Totale dataset en indeling in deelsets	9
2.3.2	Samenvattend overzicht uitgevoerde analyses	10
3	RESULTATEN	13
3.1	Algemene bevindingen	13
3.2	Vergelijking van milieu-indicatiewaarden	13
3.3	Milieu-voorspelling op basis van soorten	16
3.4	Vergelijking rekenmethoden voor bepalen indicatiewaarden	19
3.5	Alternatieve methode voor bepalen milieucondities uit soorten	21
4	SYNTHESE	24
4.1	Discussie en conclusies	24
4.2	Aanbevelingen	26
	LITERATUUR	28
	Bijlagen (als los pdf-rapport, zie volgende pagina voor een overzicht)	

- 1a vergelijking optima en GG (percentielen) Limnodata met AqMaD en B&R_geselecteerde parameters
- 1b vergelijking optima en GG (vaste breedte) Limnodata met AqMaD en B&R_geselecteerde parameters
- 2a figuren vergelijking optima watertypen
- 2b figuren vergelijking optima landschapstypen
- 2c figuren vergelijking optima waterbeheerders
- 3a vergelijking optima en GG (vaste breedte) Limnodata met WN01 en WSRL
- 3b vergelijking optima en GG (vaste breedte) Limnodata sloten met WSRL, AqMaD en B&R_geselecteerde parameters
- 4a terugvoorspelling SDK ongetransformeerd
- 4b terugvoorspelling SDK LN getransformeerd
- 4c terugvoorspelling SDK abundantie LN en parameters divers getransformeerd
- 5a vergelijking gemeten voorspeld SDK en KRW orig en In_pred2
- 5b vergelijking gemeten voorspeld KRW_In_pred2_per_watertype
- 5c vergelijking gemeten voorspeld KRW_In_pred2_per_landschapstype
- 5d vergelijking gemeten voorspeld deelsets en transformatie watertypen pred_2
- 6 - toelichting
- 6 vergelijking resultaat voorspelling bij min 2 en min 10 waarnemingen getransformeerd

1

INLEIDING

1.1 AANLEIDING

De Ecologische Sleutelfactoren (ESF's) van de STOWA zijn bedoeld om de analyse van het ecologisch functioneren en de diagnose van oppervlaktewateren te structureren. Tevens bieden ze een kapstok om daarover te communiceren. STOWA streeft naar het zo spoedig mogelijk uitwerken van tools waarmee de analyse kan worden uitgevoerd. Voor de eerste drie sleutelfactoren van stilstaande wateren zijn de tools reeds verregaand uitgewerkt, voor de vierde sleutelfactor (habitatgeschiktheid) is dit niet het geval.

Bij "habitatgeschiktheid" draait het om het geschikt zijn van de standplaats / het leefgebied voor specifieke soorten. In KRW-termen gaat het om de soorten die in de referentiesituatie of in de goede toestand mogen worden verwacht. Niet alleen is het van belang om te weten wat de habitateisen zijn van de gewenste soorten, ook is het van belang om inzicht te hebben in wat de aanwezige soorten indiceren. Hiervoor is kennis nodig van de relatie tussen het voorkomen van soorten en het habitat. De term habitat wordt vaak opgevat als de standplaats of het fysieke milieu, maar kan ook breder worden opgevat als de abiotische en biotische condities (variërend van waterkwaliteit en waterdiepte tot visstand en vegetatiebedekking).

Op dit moment wordt via verschillende sporen gewerkt aan het invullen van ESF4, de habitatgeschiktheid. Voor wat betreft de relatie tussen soorten en het milieu, wordt daarbij gezocht naar methoden die hiervoor het best geschikt zijn. Om overzicht te krijgen van de bestaande en beschikbare methoden en om een keuze hieruit te kunnen maken, wordt door STOWA een literatuurstudie uitgevoerd. Parallel hieraan wordt een vergelijkende analyse gemaakt van bestaande indicatiewaarden en indicatiewaarden die zijn afgeleid op basis van nieuwe data. STOWA heeft Nico Jaarsma gevraagd om deze vergelijkende analyse uit te voeren voor de soortgroep macrofyten (water- en oeverplanten). Voorliggend rapport is het resultaat hiervan.

1.2 HABITATEISEN VAN MACROFYTEN

Planten hebben licht, water en nutriënten nodig om te kunnen groeien. De meeste landplanten hebben een (bodem)substraat nodig om in te wortelen. Hier halen ze het grootste deel van het water en de nutriënten uit. Licht kan beperkend zijn, bijvoorbeeld in geval van landplanten door beschaduwning door andere planten.

Waterplanten hebben gewoonlijk geen gebrek aan water, licht is eerder een probleem. In troebel water of op grotere diepte is licht vaak beperkend voor de groei. In tegenstelling tot op het land, is het daarom in het water niet ongebruikelijk dat er over grote oppervlakten geen (hogere) planten groeien. Bepaalde soorten hebben aanpassingen om onder slechtere lichtcondities te groeien, bijvoorbeeld wortelende planten met bodemreserves die naar het licht toegroeien, of drijfbladplanten. Ook het substraat kan beperkend zijn, omdat het te slap is of te toxisch. Ook hier hebben bepaalde planten weer aanpassingen voor (zweven, drijven,

of zuurstof inbrengen in de bodem). Waterplanten kunnen nutriënten rechtstreeks uit de bodem of uit het water halen. Dat eerste geldt uiteraard niet voor volledig drijvende planten, zoals kroos. Naast P en N is C (koolstof) een belangrijke nutriënt. Planten kunnen dit uit de bodem, uit het water of uit de lucht halen. De vorm waarin koolstof in het water beschikbaar is, hangt af van de pH (zuurgraad). Er zijn maar weinig soorten aangepast aan wateren met een zeer lage (<6) of een zeer hoge (>8) pH.

Verschillende soorten planten hebben verschillende strategieën om met de beschikbare nutriënten, ruimte en licht om te gaan. Ze beïnvloeden daarbij ook weer hun eigen milieu. Om die reden worden planten onder verschillende milieucondities aangetroffen, dit maakt ze in potentie geschikt als indicatoren voor die milieucondities.

1.3 CENTRALE VRAAG EN DOELSTELLING

Deze studie gaat over de vraag of en hoe vegetatiegegevens kunnen worden gebruikt om iets (kwantitatiefs) te zeggen over het milieu. Kan bijvoorbeeld de zuurgraad of de fosfaatrijkdom voorspeld worden op basis van een lijstje plantensoorten? Belangrijke vragen daarbij zijn:

- Kan dit? Zo ja, hoe goed is de voorspelling?
- Welke rekenmethoden zijn dan het best bruikbaar?
- Moeten soorten- en milieudata nog worden getransformeerd?
- Op welk niveau werken indicatiewaarden het best? Landelijk, per watertype, per regio?

Het project heeft ten doel om inzichtelijk te maken welke keuzes gemaakt moeten worden om te komen tot een geschikte set aan milieu-indicatiewaarden van macrofyten.

1.4 DATASETS EN UITGEVOERDE ANALYSES

Om deze vragen te beantwoorden zijn verschillende datasets verzameld. Dit zijn bestaande indicatiewaarden (AqMaD en Bloemendaal & Roelofs) en indicatiewaarden voor macrofyten die zijn afgeleid voor Waternet (WN) en Waterschap Rivierenland (WSRL). Daarnaast zijn milieu-indicatiewaarden berekend uit gegevens over macrofyten en milieu uit de landelijke dataset van Limnodata.

Met de dataset zijn de volgende analyses uitgevoerd:

- A Vergelijking van milieu-indicatiewaarden:
 - 1 Limnodata (landelijke dataset) met AqMaD en Bloemendaal & Roelofs
 - 2 Limnodata (landelijke dataset) met deelsets hiervan (watertype, landschapsregio of waterbeheerder)
 - 3 Limnodata (landelijke dataset) met data Waternet en Waterschap Rivierenland.
- B Analyse van milieu-voorspelling op basis van soorten:
 - 1 Gemeten – voorspeld totale dataset
 - 2 Gemeten – voorspeld deelsets (watertype)
 - 3 Effect van transformatie (LN, wortel, geen)
 - 4 Methoden (regressie op gemiddelde optima of gewogen gemiddelde optima)

1.5 LEESWIJZER

Hoofdstuk 2 gaat in op de gebruikte datasets en voorbewerking, het berekenen van de indicatiewaarden en de voorspelling van milieuecondities op basis van de berekende milieu-indicatiewaarden. Deze stappen leveren input voor de onderlinge vergelijking van indicatiewaarden tussen de bestaande datasets (AqMaD en Bloemendaal & Roelofs), de landelijke (Limnodata) dataset en deelsets hiervan.

2

VOORBEWERKING EN ANALYSE

2.1 VERZAMELEN EN VOORBEWERKEN DATA

Voor de vergelijking van milieu-indicatiewaarden is gebruik gemaakt van bestaande indicatiewaarden en zijn indicatiewaarden afgeleid van datasets over het voorkomen van soorten en bijbehorende milieufactoren. Gekozen is voor de soortgroep macrofyten, omdat hiervan relatief veel informatie beschikbaar is en het gaat om een overzichtelijk aantal soorten. Voor de vergelijking zijn de volgende data gebruikt (Tabel 2.5):

- 1 Bestaande indicatiewaarden – landelijk (De Lyon & Roelofs, 1986; dataset AqMaD versie 1.0);
- 2 Bestaande indicatiewaarden – regionaal (deze in recente projecten reeds afgeleid voor WSRL, Waternet);
- 3 Nieuwe indicatiewaarden die in dit project zijn afgeleid uit de Limnodata (per watertype, landschapstype, waterbeheerder).

2.1.1 BESTAANDE INDICATIEWAARDEN

De gebruikte landelijke datasets met indicatiewaarden voor vegetatie zijn die van De Lyon & Roelofs (1986) en de indicatiewaarden uit het programma AqMaD Versie 1 (Riegman & Starink, 2009). Bij deze datasets is bij de berekening van indicatiewaarden geen onderscheid gemaakt in regio's of watertypen. De wijze waarop de indicatiewaarden zijn bepaald verschilt: Riegman & Starink maken gebruik van gemiddelden en standaarddeviatie, De Lyon & Roelofs van gewogen gemiddelden (zie Jaarsma, *et. al.*, 2015, voor een toelichting). De milieu-indicatiewaarden van ICHORS (Barendregt, *et. al.*, 1990) waren niet in een bruikbare vorm beschikbaar en zijn dus niet in de analyse meegenomen.

De indicatiewaarden van Bloemendaal & Roelofs en Riegman & Starink zijn voor zover relevant omgezet naar dezelfde eenheden (bijvoorbeeld van mmol/l naar mg/l) als die voor de betreffende parameters in de Limnodata worden gehanteerd. Soortnamen zijn gecontroleerd en uniform gemaakt (gelijk) aan Limnodata/TWN. Parameternamen zijn uniform gemaakt met Limnodata.

De beschikbare dataset van Bloemendaal & Roelofs bestaat uit indicatiewaarden voor 116 taxa en 37 parameters in oppervlaktewater, porievocht (bodem) en bodem.

De beschikbare dataset van Riegman & Starink uit indicatiewaarden voor 764 taxa en 33 parameters in het oppervlaktewater.

Regionale datasets met indicatiewaarden zijn recent afgeleid voor waterschap Rivierenland (Jaarsma, *et. al.* 2015) en voor Waternet (Jaarsma en Van den Berg, niet gepubliceerd). Deze indicatiewaarden (optima en toleranties, gewogen gemiddelden) zijn afgeleid voor waterkwaliteitsparameters (waterschap Rivierenland en Waternet) en voor Waternet aangevuld met parameters m.b.t. bodemkwaliteit, bodemsamenstelling, (hydro)morfologie etc.. Hierbij is onderscheid gemaakt in watertype (met name sloten en kanalen voor Waterschap Rivierenland en plassen voor Waternet)

2.1.2 AFLEIDEN INDICATIEWAARDEN UIT DE LIMNODATA

De gegevens van Limnodata Neerlandica zijn gebruikt voor het opstellen van milieupreferenties voor verschillende watertypen, landschapsecologische regio's en waterbeheerders (Tabel 2.5). Bij aanvang van het project zijn deze gegevens aangeleverd door Royal Haskoning-DHV. De data bestrijken de periode 1980-2012 en zijn aangeleverd als losse csv-bestanden voor de biologische groepen, een bestand met fysische-chemische data en een bestand met veldkenmerken. Hieruit zijn, na voorbewerking, indicatiewaarden afgeleid van verschillende selecties van de data en met verschillende methoden.

Hierover het volgende (Tabel 2.5):

- 1 Voor de afleiding van milieu-indicatiewaarden zijn twee methoden gebruikt: weighted averaging (berekening van optima en toleranties volgens methode ter Braak) en gewogen gemiddelden (volgens methode Bloemendaal & Roelofs), zie voor een toelichting ter Braak (1996), Bloemendaal & Roelofs (1988) en Jaarsma *et. al.*, (2015);
- 2 De indicatiewaarden zijn berekend voor de gehele landelijke dataset (macrofyten van stilstaande wateren) en voor deelsets daarvan. Daarbij is onderscheid gemaakt tussen: landschapsecologische regio's volgens OBN (laagveen, zeeklei, rivierengebied, hogere zandgronden) en hoofdwatertypen (sloten, kanalen, meren en plassen, diep en ondiep). Ook is nog onderscheid gemaakt in beheergebieden van (enkele grotere) waterbeheerders (HHNK, WHD, WSRL en WSS).

Hieronder wordt kort toegelicht welke voorbewerkingen hebben plaatsgevonden.

DATASET FYSISCH-CHEMIE LIMNODATA EN VOORBEWERKING

De fysisch-chemische dataset uit de Limnodata bevat circa 8 miljoen records (waarnemingen), van 48 parameters, op ruim 13.000 meetpunten in de periode 1980-2012. De data is aangeleverd als csv-bestand. Om hier mee te kunnen werken is het nodig om de data op te nemen in een database, te controleren op fouten/sterk afwijkende waarden en om uit de waarnemingen die op verschillende momenten zijn gedaan één zinvolle waarde te bepalen, die kan worden gekoppeld aan een vegetatie-opname. Gekozen is voor het zomergemiddelde (april-september) per jaar. In hoofdlijnen zijn de volgende voorbewerking uitgevoerd voor de fysische -chemie:

- 1 Importeren data in Access vanuit CSV (in 2 delen, 8 miljoen records!!);
- 2 Parameternamen gecontroleerd, deze zijn uniform en ongewijzigd aangehouden;
- 3 Controle op afwijkende waarden en bepalen percentiel-waarden per parameter (in MS-Access en MS-Excel);
- 4 Selectie van data tussen 1-percentiel en 99-percentiel (Access);
- 5 Berekenen zomergemiddelde waarden uit data 1-99-P per jaar en locatie (Access).

DATASET MACROFYTEN LIMNODATA EN VOORBEWERKING

De totale dataset van macrofyten in de Limnodata betreft ruim 700.000 waarnemingen, in ruim 70.000 opnamen, op ruim 28.000 locaties van circa 1.700 (niet gecontroleerde) taxa. In de database zitten veel taxa die onder verschillende namen in de database zijn opgenomen. Er zitten ook veel waarnemingen in van soorten die geen duidelijke relatie hebben met het oppervlaktewater. Er zitten verschillende opname-schalen in, die afgestemd moeten worden. Ook zijn veel opnamen niet direct te koppelen aan fysisch-chemische data. Kortom, het is nodig om de dataset op te schonen en af te stemmen op de fysisch-chemische dataset.

Daarvoor:

- 1 Zijn uit de totale soortenlijst, de soorten uit de KRW-maatlatten geselecteerd en is de naamgeving afgestemd op de TWN-lijst. De KRW-soorten zijn ten behoeve van de analyse ingedeeld in groeivormen (submers, drijfblad, kroos, emergent en overig);
- 2 Zijn de verschillende abundantieschalen, voor zover duidelijk vastgelegd, omgezet naar bedekkingspercentages;
- 3 Zijn de waarnemingen op basis van de locatiecode gekoppeld aan de fysisch-chemische dataset;
- 4 Is per locatie, de meest recente waarneming van de macrofyten (meest recente opname) geselecteerd. Hierdoor komt iedere locatie slechts eenmaal in de dataset voor;
- 5 Is het dichtstbijzijnde jaar geselecteerd waarvan ook fysisch-chemische data beschikbaar waren, daarbij is een periode van 3 jaar voor tot 1 jaar na de opname gekozen. Fysisch-chemische data van 1 jaar vóór de macrofytenopname gaan daarbij vóór data die 1 jaar ná de opnamen zijn verzameld;
- 6 Zijn de locaties gekoppeld aan het landschapstype in GIS (<http://landschapinnl.nl/bronnen-en-kaarten/cultgis>) en zijn uit de Limnodata het KRW-watertype en de waterbeheerder gehaald.

TRANSFORMATIE VAN SOORTEN- EN MILIEUDATA

Bij de analyses is in bepaalde gevallen de data getransformeerd, voordat er mee gerekend is. Dit wordt vaak gedaan, om het effect van enkele hoge waarden op de uitkomsten van de analyse te beperken. Voor soorten en milieudata is gebruik gemaakt van $\ln(x+1)$ en worteltransformatie. In hoofdstuk 3 wordt daar verder op ingegaan.

2.2 STURENDE FACTOREN VEGETATIE EN SELECTIE PARAMETERS VOOR ANALYSE

In deze stap is onderscheid gemaakt in de mate waarin factoren direct en indirect sturend zijn voor de vegetatie en is gekeken naar de beschikbaarheid van data en overlap in parameters tussen de datasets. Op basis hiervan zijn de parameters geselecteerd voor nadere analyse.

2.2.1 STURENDE FACTOREN

Runhaar & Witte (2010) maken onderscheid in factoren die de vegetatie direct en indirect beïnvloeden (operationeel versus conditioneel). Tabel 2.1 geeft enkele voorbeelden. De auteurs stellen dat de vegetatie gebruikt kan worden als indicator voor de standplaatscondities, maar dat dit alleen geldt voor de direct op de vegetatie inwerkende (operationele) factoren. Gebruik van de vegetatie om conditionerende (indirect sturende) factoren te indiceren is riskant, omdat de gevonden verbanden correlatief zijn. Anders gezegd, de vegetatie kan zeer indicatief zijn voor de beschikbaarheid van nutriënten, maar tegelijkertijd een matige indicator zijn voor het totaal P-gehalte in water of bodem. Er is namelijk een verschil tussen het totaal P-gehalte en de beschikbaarheid van P voor de vegetatie, bij dat laatste speelt bijvoorbeeld ook vorm waarin het P aanwezig is een rol en of P of iets anders limiterend is. Dit is uiteraard waar en maakt milieu-indicatie op basis van soorten complex. Anderzijds is bijvoorbeeld het (opgelost) P-gehalte in het water ook een goede indicator voor de beschikbaarheid van P. Vaak worden er ook wel duidelijke en logische verbanden gevonden met het voorkomen van soorten of groeivormen. Ook zullen er wellicht verschillen zijn tussen terrestrische en aquatische systemen; in wat er gemeten wordt en wat de indicatiewaarde van de vegetatie is voor die parameters. Daarnaast is het onderscheid tussen conditionele en operationele factoren mogelijk niet altijd helder te maken.

TABEL 2.1 OPERATIONELE EN CONDITIONELE FACTOREN VOOR DE VEGETATIE (RUNHAAR & WITTE, 2010).

factor	toelichting	voorbeelden
operationeel	werkt direct in op de plantengroei	licht, beschikbaarheid van nutriënten, maaibeheer
conditioneel	werkt indirect in op de plantengroei	totaal-gehalten van Ca, N en P, o.a. via pH en beschikbaarheid

Verdonschot (2015) heeft een overzicht gemaakt van sleutelfactoren (operationele factoren) voor organismengroepen in aquatische systemen. Daarbij is een weging gegeven voor de mate waarin die factoren worden geacht sturend te zijn per watertype. Tabel 2.2 laat dit zien voor de macrofyten.

TABEL 2.2 OVERZICHT VAN ALLE SLEUTELFACTOREN EN WEGINGSFACTOR PER KRW ORGANISMENGROEP EN HOOFDWATERTYPE, WAARBIJ 1=STERK, 10=ZWAK (VERDONSCHOT, 2015)

KRW organismengroep	Hoofd-watertype	tem-pera-tuur	licht	droog-val	stro-ming	connec-tiviteit	diepte	opper-vlak	sub-straat	zuur-graad	silicium	macro-ionen
macrofyten	ss beken	5	1	3	1				1			
	sloten	7	1	5					2-3	2		4
	ondiepe meren	7	1-2	2-5					2-3	1-2		8
	diepe meren	8	1						2-3	1-2		4
	(zwak) brak	8	1	3					4			5
		organisch materiaal	ammo-nium	zuur-stof	voedings-stoffen	ammoniak, nitriet, sulfide		macro-fyten		milieu-vreemde stoffen		direct menselijk ingrijpen
macrofyten	ss beken					2-3						
	sloten					1				1-2		4
	ondiepe meren					1-2						
	diepe meren					1-2						
	(zwak) brak					7 (t-N)						

Uit dit overzicht komt naar voren dat de hiërarchie in sturende factoren voor de macrofyten, afhangt van het watertype. Voor de stilstaande zoete wateren (sloten, ondiepe en diepe meren) is het beeld echter vrij eenduidig, hier zijn de factoren licht (1), voedingsstoffen en zuurgraad (1-2) en substraat (2-3) de belangrijkste. Verder zijn voor sloten en diepe meren de macro-ionen (4) van belang en is voor sloten direct menselijk ingrijpen (4), ofwel het (maai) beheer van belang. Opvallend is dat voor sloten milieuvreemde stoffen (1-2) als belangrijke sleutelfactor (operationele factor) worden genoemd. In het rapport wordt hierover gezegd dat milieuvreemde stoffen allesoverheersend zijn wanneer ze in zeer hoge (acuut toxische) concentraties aanwezig zijn. Dit wordt echter niet direct als groot probleem herkend uit de praktijk. Ten slotte is droogval nog een factor voor sloten (5) en ondiepe meren (2-5).

Ook in Bloemendaal & Roelofs (1988) is veel kennis te vinden over sturende factoren voor de vegetatie. Daaruit wordt onder andere duidelijk dat het, om de vegetatie te begrijpen, nodig is om naar zowel water als de waterbodem te kijken. Zo kunnen bepaalde soorten ondanks de lage nutriëntengehalten in het water, toch prima groeien omdat er voldoende beschikbaar is in de bodem. Bepaalde soorten zijn volgens hen zelfs kenmerkend voor wateren met een nutriëntenarme waterlaag en een (matig) nutriëntenrijke bodem. Dit geldt niet alleen voor de nutriënten P en N maar vooral ook voor anorganisch koolstof (CO_2 en HCO_3^-). Koolstof wordt daar zelfs de “vergeten nutriënt” genoemd (Bloemendaal & Roelofs, 1988), in dat opzicht is er nog weinig veranderd ten opzichte van enkele decennia geleden. Vaak wordt door waterbeheerders en ecologen vooral naar P en N gekeken, terwijl er toch signalen zijn dat ook in de Nederlandse wateren de beschikbaarheid van C beperkend kan zijn voor de

productiviteit als voor het voorkomen van bepaalde soorten. De beschikbaarheid van koolstof hangt in belangrijke mate samen met de zuurgraad (koolzuur-bicarbonaat evenwicht) en daarmee met de mate van buffering. Productie en afbraak beïnvloeden de zuurgraad, met de aanvoer van water kunnen koolstof en bufferende stoffen worden aangevoerd.

In tabel 2.3 is bovenstaande samengevat in een hiërarchisch overzicht van sturende factoren, waarbij een koppeling is gemaakt met de Ecologische Sleutel Factoren. In de laatste kolom zijn de factoren uitgedrukt in concrete parameters.

Licht is de eerste factor en zuurgraad staat samen met buffering op de tweede plaats (zie Bloemendaal & Roelofs, 1988), gevolgd door nutriënten, substraat, macro-ionen en toxiciteit (voor de afbraakproducten ammonium en sulfide). De hier gepresenteerde hiërarchie is logisch en onderbouwd door de literatuur. Maar het is daarbij van belang om te bedenken dat er een duidelijke samenhang is tussen de stuurfactoren, en dat ze dus niet los van elkaar kunnen worden gezien. In het water heeft het doorzicht vaak een relatie met de nutriëntenbelasting en het nutriëntengehalte, hoewel dit voor verschillende watertypen (sloten, ondiepe en diepe meren) weer anders uitpakt. Het lichtklimaat is ook meer dan alleen het doorzicht, wat weer niet los kan worden gezien van de waterdiepte. Zo zijn er veel dwarsverbanden en relaties, de uitdaging is deze uiteen te rafelen.

In dit rapport worden meestal mono-factoriele relaties tussen abiotische parameters en vegetatie onderzocht, soms multifactorieel (zoals de ratio doorzicht/diepte). In werkelijkheid zijn de relaties waarschijnlijk multifactorieel (hier wordt in de discussie en aanbevelingen nader op ingegaan).

TABEL 2.3 HIËRARCHISCH OVERZICHT VAN STURENDE FACTOREN VOOR DE VEGETATIE VAN STILSTAANDE ZOETE OPPERVLAKTEWATEREN EN KOPPELING AAN ECOLOGISCHE SLEUTEL FACTOREN.

ESF	factor(en)	toelichting	parameters
2	licht	Planten hebben licht nodig voor fotosynthese, in oppervlaktewater is licht al snel beperkend. Er is o.a. een relatie tussen de beschikbaarheid van licht en het voorkomen van groeivormen: van laag blijvende wortelende soorten in zeer helder water, via soorten die naar het licht toe groeien of zweven naar drijfbladplanten bij een geringe lichtbeschikbaarheid.	doorzicht, extinctie en ratio van doorzicht/diepte
4	zuurgraad en buffering	Dit is eigenlijk een complex van factoren die samenhangt met zuurgraad en buffering en de beschikbaarheid van koolstof. Deze factoren zijn sterke mate sturend voor het voorkomen van de vegetatie, in zure en zwak gebufferde milieus kunnen zich maar een beperkt aantal soorten handhaven, hetzelfde geldt ook voor sterk basische milieus	pH, alkaliniteit, (bi) carbonaat
1+3	nutriënten	Planten hebben nutriënten nodig om te groeien, de meest bekende zijn fosfor en stikstof. Soorten kunnen deze uit de bodem of uit het water halen.	beschikbare en totale fracties van P en N in water en waterbodem
3+4	substraat	Het substraat (de sliblaag en de onderliggende bodem) is van belang voor worteling en speelt een belangrijke rol in allerlei biogeochemische processen zoals afbraak van organisch materiaal en binding en nalevering van nutriënten. Het substraat verschilt o.a. in de samenstelling, dikte en consistentie van de sliblaag en de onderliggende bodem. Op minerale (zand of klei) bodems worden vaak andere soorten aangetroffen dan op organische (vast of afgebroken veen of detritus) bodems.	organisch stofgehalte, dikte sliblaag, vochtgehalte slib
4	macro-ionen	De zuurgraad en mate van buffering zijn in het veld vaak duidelijk aan elkaar gecorreleerd (een lage pH gaat vaak samen met lage concentraties aan macro-ionen). Het gaat vooral om chloride en sulfaat (ook overlap met toxiciteit) en om calcium, magnesium, natrium en kalium.	Cl, SO ₄ (Ca, Mg, Na, K)
3+8	toxiciteit	Voor bepaalde soorten zijn ammonium/ammoniak en sulfides toxisch. Dit hangt sterk samen met afbraakprocessen in de bodem en de beschikbaarheid van ijzer (Fe) en zwavel (S). Er is nog vrij weinig bekend over de mate waarin dit echt een rol speelt, naar verwachting geldt dit vooral voor zeer voedselrijke en/of hoog belaste wateren	NH ₄ en sulfide (S ²⁻) in bodemvocht

2.2.2 SELECTIE PARAMETERS VOOR NADERE ANALYSE

Tabel 2.3 laat zien welke factoren als (direct) sturend worden gezien voor de vegetatie. In de tabel is ook aangegeven welke parameters hierop betrekking hebben. Daarbij is (in de lijn van Runhaar en Witte, 2000) de vraag in hoeverre deze parameters de factoren voldoende goed representeren en dus een oorzakelijk verband kunnen indiceren. Besloten is deze parameters zoveel mogelijk in de analyse mee te nemen en per factor(complex) achteraf te kijken welke parameter het duidelijkste verband laat zien met de soorten. Soms is er weinig keuze (niet of slechts één parameter goed gemeten), soms kunnen meerdere parameters worden vergeleken. Uiteindelijk moet het resultaat met gezond verstand worden geïnterpreteerd.

Binnen de beschikbare datasets zijn veel van de genoemde parameters in het oppervlaktewater wel gemeten, dat geldt echter niet voor de bodem gerelateerde parameters (de laatstgenoemde parameters zijn alleen in dataset Waternet opgenomen). De database van de Limnodata is in belangrijke mate bepalend voor de vergelijking, hieruit worden indicatiewaarden van soorten bepaald voor verschillende indelingen (watertype, regio, beheerder) van de dataset. Voor de vergelijking zijn daarom de volgende parameters relevant:

- Lichtklimaat: Doorzicht (diepte zit maar zeer beperkt in de Limnodata);
- Zuurgraad en buffering: pH, alkaliniteit, (bi)carbonaat;
- Nutriënten: P, PO₄, N, NO₃, NH₄ in de waterlaag;
- substraat: geen bruikbare data beschikbaar;
- macro-ionen: Cl, SO₄, Ca, Mg, Na, Ka;
- toxiciteit: van bodemvocht zijn geen data beschikbaar.

Bij de analyse voor Waternet is de relatie tussen de vegetatie en de bodemparameters en de ratio doorzicht/diepte wel onderzocht; hier wordt in de discussie nader op ingegaan.

2.3 OVERZICHT DATASETS EN UITGEVOERDE ANALYSES

2.3.1 TOTALE DATASET EN INDELING IN DEELSETS

De waarnemingen uit de Limnodata zijn (na voorbewerking en selectie) ingedeeld in een trainingsset (90% van de waarnemingen) en een test-set (10%). De trainingsset is gebruikt om de milieu-indicatiewaarden mee af te leiden. De test-set om het resultaat met een onafhankelijke set van waarnemingen te valideren. Verder is de set opgedeeld in deelsets, die gebruikt zijn bij het bepalen van de indicatiewaarde van soorten binnen een bepaald watertype, een bepaalde landschapsregio of voor specifieke waterbeheerders (Tabel 2.4 en 2.5). De gegevens uit Limnodata zijn daarbij uitgesplitst per watertype of landschapsecologische regio, met andere woorden: de wateren per landschapsecologische regio zijn niet verder onderverdeeld in watertypen en *vice versa*.

TABEL 2.4 AANTAL UNIEKE WAARNEMINGEN (MONSTERS) PER DEELSET EN VERDELING OVER TRAINING- EN TESTSET.

aantal monsters			aantal monsters		
waterbeheerder	training	test	hoofdtype	training	test
AGV	180	22	brakke meren	924	103
HDSR	102	5	diepe meren	153	13
HHD	295	30	grote diepe kanalen	86	9
HHNK	757	78	grote ondiepe kanalen	207	22
HHR	296	27	ondiepe meren	835	83
HHSK	191	33	sloten	1718	166
KUN	9	0	vaarten en kanalen	847	93
PGR	10	1	-	528	65
PNH	36	6	Eindtotaal	5298	554
POV	63	10			
PRF	40	3			
PRU	56	4			
WAM	57	4			
WBD	167	14			
WF	439	48			
WGS	154	15			
WHA	73	4			
WHD	483	42			
WNZV	49	2			
WPM	38	1			
WRD	138	19			
WRIJ	37	4			
WRO	33	3			
WRW	93	13			
WSD	1	0			
WSRL	336	35			
WSS	651	64			
WSV	94	11			
WVE	71	7			
WVV	91	17			
WZZ	258	32			
Eindtotaal	5298	554			

aantal samples		
landschapsregio	training	test
Droogmakerijen	657	79
Heuvelland	4	0
Kustzone	233	22
Laagveengebied	851	97
Rivierengebied	526	46
Veenkoloniën	120	18
Verstedelijk gebied	559	58
Zandgebied	811	76
Zeekleigebied	1524	156
-	13	2
Eindtotaal	5298	554

2.3.2 SAMENVATTEND OVERZICHT UITGEVOERDE ANALYSES

Tabel 2.5. geeft een overzicht van de gebruikte datasets, en de analyses waarvoor deze gebruikt zijn. Hieronder worden de uitgevoerde analyses kort toegelicht.

VERGELIJKING MILIEU-INDICATIEWAARDEN

Voor de vergelijking van milieu-indicatiewaarden zijn de indicatiewaarden uit de verschillende datasets en deelsets in grafieken (scatterplots) tegen elkaar uitgezet. Hiermee kan tegelijkertijd worden beoordeeld of de indicatiewaarden in absolute zin vergelijkbaar zijn (vergelijking met de 1:1 lijn) en of ze een vergelijkbaar patroon laten zien (trendlijn met een correlatiecoëfficiënt (r^2)).

VOORSPELLING MILIEU OP BASIS VAN BEREKENDE MILIEU-INDICATIEWAARDEN

Met de berekende milieu-indicatiewaarden van de soorten, kan een voorspelling worden

gedaan van de milieucondities op een locatie waar een vegetatie-opname is uitgevoerd. Hier zijn door diverse auteurs (o.a. Ter Braak, 1996, Birks, *et. al.*, 1990) methoden ontwikkeld. De voorspelling is hier gedaan door “weighted averaging” (berekening van de optima, ofwel de naar abundantie gewogen gemiddelden) in combinatie met “linear deshrinking” (lineaire regressie op de berekende gewogen gemiddelden). Dit laatste is nodig om het effect van het “krimpen” van de gradiënt door het tweemaal gewogen middelen tegen te gaan. Per (deel)set en parameter wordt hiervoor een lineair regressiemodel afgeleid. Het regressiemodel wordt gefit op de oorspronkelijke data (trainingsset, waarbij X = gewogen optimum en Y = oorspronkelijke meetwaarde). Dit levert een regressievergelijking op, die kan worden gebruikt voor nieuwe data. Voor het evalueren van de voorspelling van milieucondities uit de berekende milieu-indicatiewaarden is gebruik gemaakt van een onafhankelijke dataset (random selectie van 10% van de waarnemingen).

EFFECT VAN TRANSFORMATIE

Bij de voorspelling van het milieu uit de soorten is gekeken naar het effect van de keuze voor transformatie van soort-abundanties en waarden van milieuvariabelen. Deze transformatie wordt vooraf (aan het bepalen van de optima en het regressiemodel) gedaan, achteraf worden de voorspelde waarden weer teruggetransformeerd. Er is een vergelijking gemaakt van de volgende transformaties: geen, $\ln(x+1)$ en wortel.

TABEL 2.5 SAMENVATTEND OVERZICHT UITGEVOERDE ANALYSES.

		aantal samples		analyse		
(deel)set		training	test	vergelijking optima	voorspelling milieu	effect transformatie
Bestaande datasets						
-	Bloemendaal & Roelofs			x		
-	AqMaD			x		
-	WSRL			x		
-	Waternet			x		
Limnodata						
-	totaal	5298	554	x	x	x
Limnodata watertype						
-	brakke meren	924	103	x	x	x
-	diepe meren	153	13	x	x	x
-	grote diepe kanalen	86	9	x	x	x
-	grote ondiepe kanalen	207	22	x	x	x
-	ondiepe meren	835	83	x	x	x
-	sloten	1718	166	x	x	x
-	vaarten en kanalen	847	93	x	x	x
Limnodata regio						
-	Droogmakerijen	657	79	x		
-	Kustzone	233	22	x		
-	Laagveengebied	851	97	x		
-	Rivierengebied	526	46	x		
-	Veenkoloniën	120	18	x		
-	Verstedelijkt gebied	559	58	x		
-	Zandgebied	811	76	x		
-	Zeekleigebied	1524	156	x		
Limnodata waterbeheerder						
-	HHNK	757	78	x		
-	WHD	483	42	x		
-	WSRL	336	35	x		
-	WSS	651	64	x		

3

RESULTATEN

De analyse heeft veel resultaten en verschillende nuttige inzichten opgeleverd. In dit hoofdstuk worden deze kort gepresenteerd, daarbij wordt verwezen naar de bijlagen waar de resultaten in uitgebreide vorm zijn opgenomen. In onderstaande paragrafen worden de resultaten besproken met betrekking tot:

- Algemene bevindingen
- Met betrekking tot de vergelijking van milieu-indicatiewaarden
- Met betrekking tot milieu-voorspelling op basis van soorten

3.1 ALGEMENE BEVINDINGEN

Het doel van deze studie is vooral om te bepalen welke indicatiewaarden en/of methoden om indicatiewaarden te berekenen het best bruikbaar zijn. Daarvoor is de soortgroep macrofyten gekozen, dit zijn de soorten uit de KRW-maatlatten, waarbij de focus heeft gelegen op soorten met de groeivormen submers, drijfblad en kroos, maar ook de emergente en overige soorten zijn betrokken. De analyses geven echter ook inzicht in de relatie tussen macrofyten en milieu, zoals die uit de diverse datasets naar voren komen. Ten aanzien daarvan kunnen enkele algemene bevindingen worden genoemd:

Relatie macrofyten milieu: Uit de analyse blijkt dat aquatische macrofyten in de Limnodata een duidelijke relatie vertonen met de milieuomstandigheden. Voor de macrofyten zijn relaties gevonden met de volgende factorcomplexen en parameters; in volgorde van belangrijkheid:

- Zuurgraad en buffering: pH, alkaliniteit, bicarbonaat en calcium;
- Zouten / waterchemie: chloride, natrium, magnesium, kalium en sulfaat;
- Nutriënten: ortho-P en totaal-P, totaal-N.

Relatie macrofyten lichtklimaat: Een duidelijke relatie met het lichtklimaat (doorzicht) is niet gevonden. Voldoende licht is een eerste vereiste voor plantengroei, uit de analyse voor Waternet kwam de ratio doorzicht/diepte heel duidelijk naar voren. In de ontvangen Limnodata-set is echter meestal alleen het doorzicht bekend, zodat de ratio zicht/diepte niet berekend kon worden....

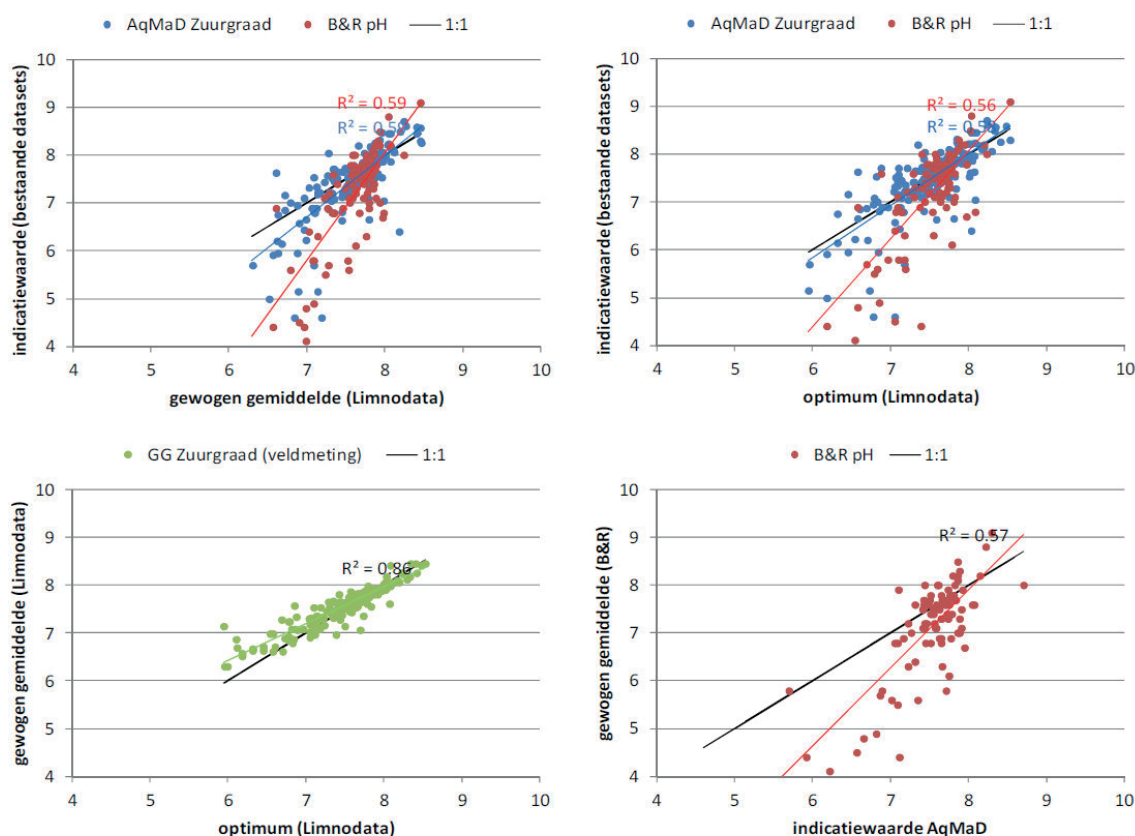
3.2 VERGELIJKING VAN MILIEU-INDICATIEWAARDEN

De onderlinge vergelijking van milieuindicatiewaarden (MIW) van de verschillende datasets maakt het mogelijk om iets te zeggen over in hoeverre de gevonden patronen robuust zijn (leveren de verschillende datasets hetzelfde beeld?). Ze geven daarmee ook inzicht in nut en noodzaak van het indelen van datasets in deelsets op basis van watertype of regio en verschillen tussen methoden. Hieronder worden de bevindingen besproken van de vergelijking van de indicatiewaarden die zijn afgeleid van de stagnante wateren in de Limnodata met achtereenvolgens: landelijke sets, deelsets gebaseerd op landschapstype en regionale sets voor verschillende waterbeheerders (zie Tabel 2.5).

Vergelijking MIW's Limnodata met bestaande landelijke sets: de vergelijking van de bestaande indicatiewaarden onderling en met de nieuw berekende landelijke indicatiewaarden (op basis van Limnodata), is opgenomen in bijlage I (Ia en Ib). De vergelijking is gedaan voor zowel de berekende optima als gewogen gemiddelden (GG) met de reeds bestaande landelijke datasets van AqMaD en Bloemendaal en Roelofs (B&R). Onderstaande figuur geeft een voorbeeld van de vergelijking voor de zuurgraad. Ter toelichting: per parameter zijn telkens 4 figuren weergegeven, in die figuren vertegenwoordigt iedere waarneming (elk symbool) één soort. Telkens wordt de landelijke milieuindicatiewaarde van die soort volgens methode_A op de x-as vergeleken met die volgens methode_B op de y-as, waarbij (van linksboven met de klok mee):

- de milieu-indicatiewaarden van AqMaD en B&R (y-as) worden vergeleken met het gewogen gemiddelde (GG methode B&R) op basis van de Limnodata;
- de milieu-indicatiewaarden van AqMaD en B&R (y-as) worden vergeleken met het optimum op basis van de Limnodata;
- de milieu-indicatiewaarden van AqMaD en B&R onderling worden vergeleken;
- het optimum en het gewogen gemiddelde (GG) op basis van de Limnodata onderling worden vergeleken.

FIGUUR 3.1 BOVEN: VERGELIJKING VAN HET GEWOGEN GEMIDDELDE (LINKS) EN HET OPTIMUM (RECHTS) VOOR DE ZUURGRAAD (VELDMETING) UIT DE LIMNODATA MET DE INDICATIEWAARDE UIT BESTAANDE DATASETS. ONDER: VERGELIJKING VAN OPTIMA MET GEWOGEN GEMIDDELDE UIT DE LIMNODATA (LINKS) EN DE INDICATIEWAARDEN VAN AQMAD MET HET GEWOGEN GEMIDDELDE VOLGENS BLOEMENDAAL&ROELOFS (RECHTS) VOOR DE ZUURGRAAD.



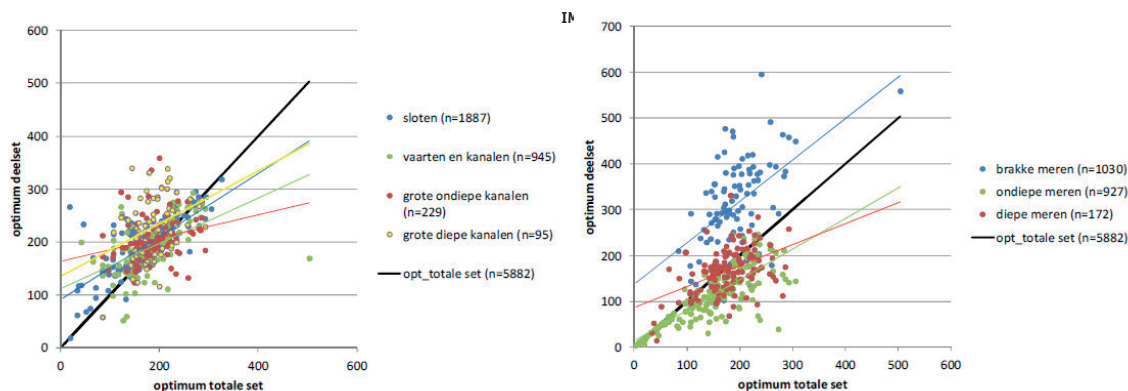
De vergelijking (zie bijlage I) laat zowel overeenkomsten als verschillen zien. In het algemeen is de trend in indicatiewaarden van soorten gelijk, soms zijn er systematische verschillen te zien (bijvoorbeeld voor calcium bij B&R). Soms is het patroon vergelijkbaar, maar zijn er grote verschillen voor individuele soorten zichtbaar; soms is het verband onduidelijk (parameter is

weinig sturend). In grote lijnen zijn de verbanden en indicatiewaarden dus wel vergelijkbaar. Dat laat onverlet dat er wel grote verschillen zijn voor individuele soorten.

NB! Overigens is het verschil tussen bijlage Ia en Ib de wijze waarop het gewogen gemiddelde (GG B&R) is berekend, in bijlage Ia bestaan de klassen uit een gelijk aantal waarnemingen, bij Ib uit een gelijke (of vooraf gedefinieerde) klassebreedte. Dit levert duidelijk andere resultaten op, in paragraaf 3.4 en bij de discussie wordt daarop ingegaan.

Vergelijking landelijke MIW's Limnodata met de deelsets: de verschillen in milieu-indicatiewaarden tussen deelsets (ingedeeld op basis van landschapstype of watertype) komen duidelijk naar voren, bijlage 2 laat dat zien (2a t/m c). In de figuren is telkens een vergelijking gemaakt van het optimum van een soort op basis van de landelijke Limnodata dataset (x-as), in vergelijking met het optimum op basis van de Limnodata deelset (y-as). Per deelset is een trendlijn weergegeven, om te laten zien of de trend in de deelset vergelijkbaar is. Er is gekozen om voor de vergelijking alleen het optimum te gebruiken. Ter illustratie hieronder de vergelijking tussen de berekende optima voor bicarbonaat, voor de totale landelijke dataset en de deelsets op basis van watertypen. De linker figuur laat de lijnvormige typen en de rechter figuur de meren en plassen en brakke wateren zien. Vooral de optima van brakke wateren wijken systematisch af (liggen ca. een factor 2 hoger), de trends zijn wel vergelijkbaar.

FIGUUR 3.2 VERGELIJKING VAN DE OPTIMA VAN DE KRW-TAXA VOOR BICARBONAAT DIE ZIJN BEREKEND VOOR ENKELE DEELSETS VAN DE LIMNODATA MET DE



Uit de vergelijking komt naar voren dat:

- de indeling in hoofdwatertypen (bijlage 2a, sloten, vaarten en kanalen, grote ondiepe kanalen, grote diepe kanalen, ondiepe meren, diepe meren, brakke wateren) laat zien dat vooral de brakke meren sterk afwijken af van de zoete (ondiepe en diepe) meren voor wat betreft ionen (o.a. bicarbonaat, Ca, Cl, SO₄) en nutriënten (met name P en PO₄). De zoete meren en plassen wijken minder af van de zoete lijnvormige wateren, het grootste verschil lijkt het optimum voor nutriënten, wat in sloten vaak duidelijk hoger ligt;
- de indeling in landschapstypes (bijlage 2b, laagveengebied, zandgebied, etc..) laat voor verschillende parameters een verschil zien, waarin de verdeling van hoog en laag-Nederland duidelijk zichtbaar is, met het rivierengebied als overgangszone;
- Dit verschil komt ook (zij het wat minder duidelijk) tot uitdrukking bij de verdeling in waterbeheerders (HHNK, WHD, WSRL, WSS).

Vergelijking landelijke MIW's Limnodata met regionale datasets: Deze vergelijking is gedaan voor de dataset van WSRL (sloten) en van Waternet (ondiepe plassen). In bijlage 3a is een vergelijking gemaakt van de optima en gewogen gemiddelden (GG B&R) van beide sets met de

milieu-indicatiewaarden die zijn bepaald op basis van de landelijke Limnodata set en zijn ze onderling vergeleken. In bijlage 3b is voor WSRL een vergelijking gemaakt met de sloten uit de Limnodata. In alle gevallen is het verband op zijn zachtst gezegd matig, in veel gevallen is er helemaal geen verband. Dit is opvallend en een belangrijke constatering. Wat precies de reden is moet nader worden bekeken, wat waarschijnlijk zeker een rol speelt is:

- De lokale (lees per deelset van landschaps- of watertype) aanwezige gradiënt in de milieufactor(en) is “kort” en omvat slechts een deel van de totale gradiënt, binnen dit traject reageren soorten niet of nauwelijks op de verschillen. Voor de respons van de vegetatie op de waterkwaliteit binnen het gebied van WSRL is dat ook al geconstateerd (Jaarsma, *et. al.*, 2015), voor de data van Waternet werden vooral andere factoren gevonden (bodemsamenstelling, doorzicht/diepte) die de variatie in soortensamenstelling konden verklaren;
- Er spelen lokale, interacterende factoren een rol. Zo is bijvoorbeeld bekend van WSRL dat de aanvoer van ijzer in de (rivier)kwel zorgt voor binding van P in/aan de bodem. Hierdoor zijn de P-gehalten in het oppervlaktewater vaak relatief laag, terwijl er in de bodem voldoende voor planten beschikbaar P aanwezig is en het water dus eigenlijk wel voedselrijk is.

Vooruitlopend op de discussie kan hieruit worden geconcludeerd dat deze bevinding **pleit voor het baseren van milieu-indicatiewaarden op grotere (landelijke) datasets met een voldoende lange gradiënt.**

3.3 MILIEU-VOORSPELLING OP BASIS VAN SOORTEN

De volgende stap is om van de berekende optima te komen tot een voorspelling van de milieu-condities. Daar zijn verschillende methoden voor, hier is gewerkt met zogenaamde “weighted averaging” gecombineerd met “linear deshrinking”. Deze methode maakt een voorspelling in 2 stappen:

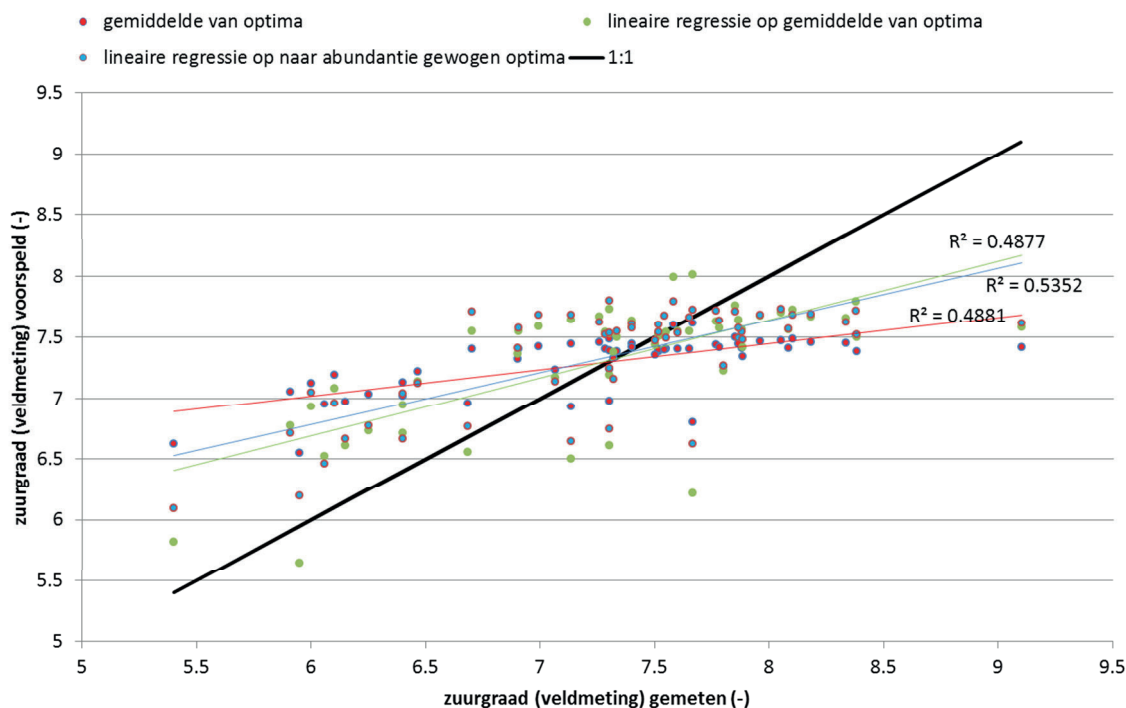
- 1 De eerste stap is het bepalen van het (gewogen) gemiddelde optimum van de aanwezige soorten per opname. Dit wordt berekend uit de optima van de aanwezige soorten, voor een specifieke milieuvariabele, al dan niet gewogen met de abundantie van die soorten. In dit geval zijn verschillende methoden onderling vergeleken, waarbij zowel met gemiddelde optima is gerekend als met naar abundantie gewogen gemiddelde optima;
- 2 Vervolgens wordt het (gewogen) gemiddelde optimum door middel van lineaire regressie “vertaald” naar een schatting van de waarde van een milieuvariabele. Dit is nodig omdat het gewogen optimum tot stand is gekomen door 2 x middelen (eerst bij berekening optima, dan bij berekening gewogen gemiddelde). Hierdoor “krimpt” het bereik van de voorspelling, wat kan worden opgelost door lineaire regressie (zie als voorbeeld figuur 3.3).

Het regressiemodel wordt gefit op de oorspronkelijke data (trainingsset, waarbij X = (gewogen) gemiddelde optimum en Y = oorspronkelijke meetwaarde). Dit levert een regressievergelijking op, die kan worden gebruikt voor nieuwe data (test-set). Er zijn bij deze procedure nog een aantal mogelijkheden:

- 1 Keuze voor transformatie van soort-abundanties en waarden van milieuv variabelen (geen, $\ln(x=1)$, wortel);
- 2 Regressie op de gemiddelde optima, de naar relatieve abundantie gewogen gemiddelde optima, de naar absolute abundantie gewogen gemiddelde optima en weging naar tolerantie. In dit geval is gekozen voor weging naar absolute abundantie.

FIGUUR 3.3

VERGELIJKING VAN GEMETEN WAARDE VAN DE ZUURGRAAD (VELDMETING) VAN DE TEST-SET OP DE X-AS MET DE VOORSPELLING VAN DE ZUURGRAAD OP BASIS VAN DE AANWEZIGE SOORTEN VOLGENS VERSCHILLENDE METHODEN: 1) HET GEMIDDELDE VAN DE OPTIMA VAN DE AANWEZIGE SOORTEN PER OPNAME, 2) NA LINEAIRE REGRESSIE OP DE GEMIDDELDE VAN DE OPTIMA VAN DE AANWEZIGE SOORTEN EN 3) NA LINEAIRE REGRESSIE OP DE NAAR ABUNDANTIE GEWOGEN GEMIDDELDE VAN DE OPTIMA VAN DE AANWEZIGE SOORTEN.



De analyse levert de volgende resultaten en inzichten op:

Prestatie modellen trainingsset + effect transformatie: bij het afleiden van de modellen op basis van de trainingsset (gehele dataset) wordt een eerste indruk verkregen van de mate waarin met de optima een goede voorspelling kan worden gedaan. In bijlage 4a t/m c zijn de resultaten van de terugvoorspelling weergegeven voor de modellen zonder transformatie, de modellen met $\ln(x+1)$ transformatie en voor de modellen gebaseerd op een $\ln(x+1)$ transformatie van abundanties en de transformatie van milieuvariabelen - (géén, $\ln(x+1)$ of wortel - met de geringste scheefheid (skewness). Dit laatste is een indicatie voor de mate waarin de data normaal verdeeld zijn (hoe lager de skewness, hoe beter normaal verdeeld).

Overall is de performance (in termen van r^2) van de modellen die zijn gebaseerd op getransformeerde data het best, met name de $\ln(x+1)$ getransformeerde data (zie tabel 3.1.). Hiervan presteren de modellen voor het factorcomplex alkaliniteit, bicarbonaat, calcium (en zuurgraad) het best. Dat wil niet direct zeggen dat dit de beste modellen zijn. Zo is de r^2 een slechte maat voor de absolute waarde van de voorspelling (gaat alleen om de helling), beter is de RMSEP (wortel van de gekwadeerde afwijkingen). Ook laten de figuren zien dat de relatie vaak niet lineair is of alleen opgaat voor een bepaald traject van de milieuvariabele. Dat wordt toegelicht aan de hand van figuur 3.4 van calcium.

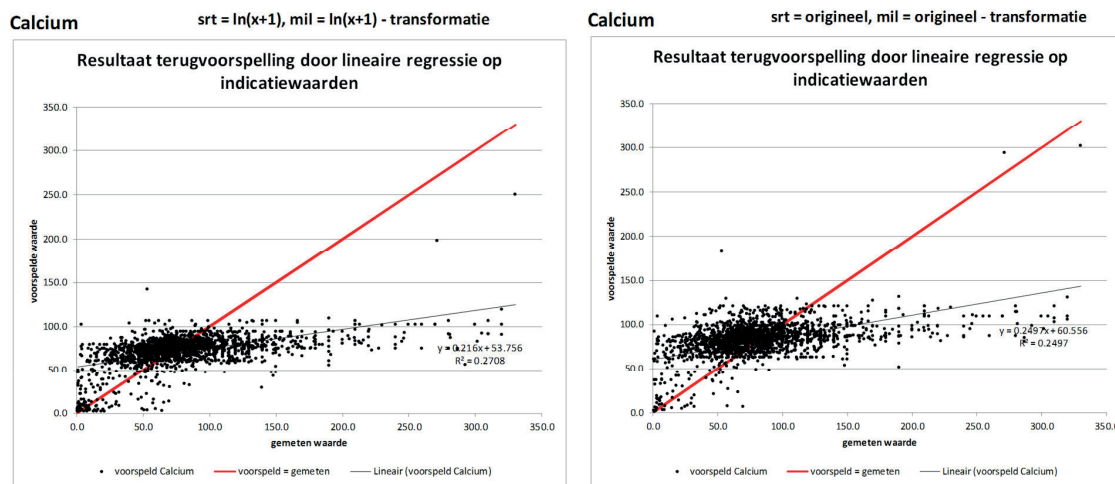
TABEL 3.1

CORRELATIECOËFFICIËNT (R^2) VOOR DE TERUGVOORSPELLING VAN MODELLEN GEBASEERD OP DE GROEVORMEN SUBMERS, DRIJFBLAD EN KROOS. ER ZIJN VERSCHILLENDE COMBINATIES TOEGEPAST VAN TRANSFORMATIE VAN DE ABUNDANTIE VAN SOORTEN ($\ln(x+1)$ OF GEËN, 1^E RIJ) EN VAN WAARDEN VAN MILIEUVARIABLEN ($\ln(x+1)$, WORTEL OF GEËN 2^E RIJ). NB! NIET ALLE COMBINATIES ZIJN DOORGEREKEN. R^2 IS GEBASEERD OP DE GETRANSFORMEERDE WAARDEN, GROEN GEMARKEERD IS HET MODEL MET DE HOOGSTE R^2 PER PARAMETER.

transformatie soorten		$\ln(x+1)$			origineel
parameter	milieu	$\ln(x+1)$	origineel	wortel	origineel
Alkaliniteit		0.62		0.61	0.56
Ammonium		0.10		0.10	0.06
Bicarbonaat		0.57	0.31		0.30
Calcium		0.50		0.37	0.25
Chloride		0.34	0.39		0.37
Doorzicht		0.16		0.16	0.15
Fosfor-totaal		0.30		0.32	0.24
Kalium		0.30			0.31
Magnesium		0.40			0.42
Natrium		0.33	0.27		0.26
Nitraat		0.07		0.07	0.03
ortho-Fosfaat		0.28		0.31	0.24
Stikstof-totaal		0.18			0.12
Sulfaat		0.28			0.24
Zuurgraad		0.33	0.32		0.28

FIGUUR 3.4

RESULTAAT VAN DE TERUGVOORSPELLING VOOR CALCIUM. LINKS, HET BEST PRESTERENDE MODEL OP BASIS VAN DE R^2 VOLGENS TABEL 3.1., RECHTS HET RESULTAAT ZONDER TRANSFORMATIE VAN SOORTEN OF MILIEU.



De figuur laat een aantal zaken zien, ten eerste dat de “bulk” van de waarnemingen van calcium ligt tussen circa 50-100 mg/l. NB! Dit heeft effect op de ligging van de optima!! Binnen dit traject is het calciumgehalte waarschijnlijk niet onderscheidend voor de vegetatie. Voor waarnemingen van calcium beneden de 50 mg/l ligt het anders, voor deze locaties worden op basis van de vegetatie ook daadwerkelijk lagere calciumgehalten voorspeld, voor waarden boven de 100 mg/l ligt de voorspelling echter weer nauwelijks hoger. Samenvattend: de vegetatie is in staat het verschil tussen wateren met een calciumgehalte van circa < 25-50 mg/l en > 50 mg/l te indiceren. Een voorspelling van het absolute calciumgehalte op basis van de soorten is echter “hachelijk”. Wel kan de voorspelling mogelijk worden verbeterd door de dataset zo te kiezen dat er een betere spreiding van waarden is over de gradiënt.

In bijlage 4 zijn nog andere voorbeelden te vinden, in veel gevallen is het beeld vergelijkbaar (o.a. pH, bicarbonaat, sulfaat, totaal-P).

Prestatie modellen test-set + effect transformatie: de resultaten van de terugvoorspelling van de trainingsset laten al voor een deel de beperkingen van de modellen zien. In bepaalde gevallen lijkt er echter toch wel een bruikbaar verband te bestaan. Echter, “the proof of the pudding is in the eating”; de onafhankelijke test-set moet duidelijk maken in hoeverre dat verband er ook echt is. Hierover het volgende:

- Bijlage 5a laat het resultaat van de voorspelling zien voor de totale dataset (met alle KRW-soorten) en voor de dataset met alleen submers, drijfblad en kroos. Daarbij is een vergelijking gemaakt van de modellen gebaseerd op de originele (niet-getransformeerde) waarden en op de getransformeerde waarden. In het algemeen geldt dat de modellen gebaseerd op de totale lijst aan KRW-soorten en op de LN-getransformeerde waarden de beste resultaten opleveren. Een probleem van de modellen gebaseerd op de originele getallen is dat ze in het lage bereik vaak negatieve waarden opleveren.
- Bijlage 5b laat voor de best presterende modellen (LN getransformeerd, totale lijst KRW soorten) zien hoe de voorspelling samenhangt met de watertypen. Daarbij zijn zowel de originele als getransformeerde waarden uitgezet. Hieruit blijkt dat de brakke wateren en de sloten vaak de grootste afwijking van de voorspelling laten zien. Vaak ziet de voorspelling er in de grafiek met de getransformeerde waarden best aardig uit, maar zijn de afwijkingen in de originele waarde soms behoorlijk groot. Bijlage 5c laat dit zien voor de landschapstypen.
- Bijlage 5d tenslotte, laat zien hoe de voorspelling uitpakt voor enkele watertypen, wanneer de optima en de modellen worden afgeleid van de deelset van dat watertype (in plaats van één grote dataset). Hieruit blijkt dat het soms wellicht wel wat betere resultaten oplevert, maar dat er in andere gevallen vrijwel geen verband is gevonden. Daardoor presteren sommige modellen voor bepaalde watertypen weer slechter dan het “totaal-model”.

3.4 VERGELIJKING REKENMETHODEN VOOR BEPALEN INDICATIEWAARDEN

Er zijn twee methoden gebruikt voor het bepalen van milieuindicatiewaarden uit de Limnodata. De eerste zijn de “optima”, dit zijn de “naar abundantie gewogen gemiddelden” volgens Ter Braak (1996). De tweede zijn de “gewogen gemiddelden” (GG) volgens de methode uit Bloemendaal & Roelofs (1988), dit zijn de “naar relatieve voorkomen in klassen gewogen gemiddelden”. Zie ook Jaarsma *et. al.*, (2015).

Het gewogen gemiddelde volgens Bloemendaal & Roelofs is duidelijk iets anders dan het optimum, de eerste hangt meer gewicht aan “afwijkende” waarnemingen. Dat is een kracht, het maakt de indicatiewaarde (GG) minder gevoelig voor de ligging van het zwaartepunt in de dataset. Dit is echter tevens een risico; wanneer er weinig waarnemingen zijn in een bepaald bereik, kan er veel gewicht komen te hangen aan één enkele waarneming. Wat echter lastiger is om mee om te gaan, is dat de indeling van de parameterwaarden in klassen in sterke mate bepalend is voor het gewogen gemiddelde volgens de methode Bloemendaal & Roelofs. Onderstaand wordt dat geïllustreerd aan de hand van twee verschillende (reproduceerbare) manieren om klassen in te delen:

- 1 Indeling van klassen door uit te gaan van een gelijke klassebreedte;
- 2 Indeling van klassen door uit te gaan van een gelijk aantal waarnemingen per klasse (dit zijn eigenlijk klassengrenzen op basis van percentielen).

Tabel 3.2. geeft dit weer voor de pH als voorbeeld. Links de klassenindeling bij gelijke klassebreedte (= 0,4 pH-eenheid) en rechts die met een gelijk aantal waarnemingen (10-P, 20-P, 30-P, etc...). Dit is in figuur 3.5. ook grafisch geïllustreerd.

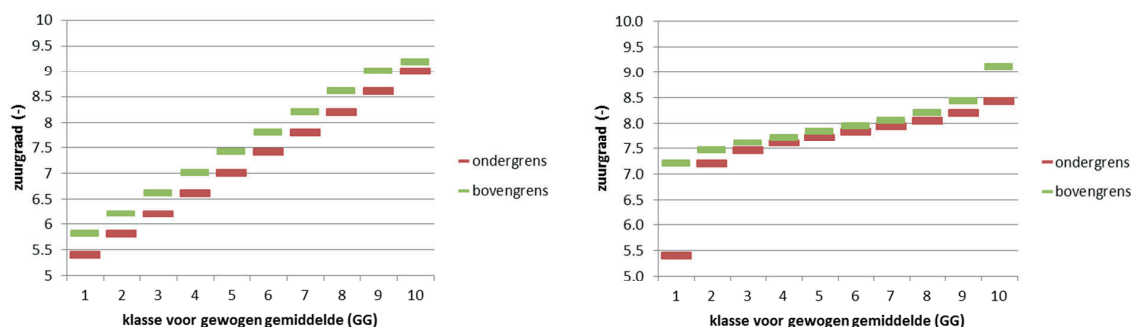
TABEL 3.2

KLASSENGRENZEN GG GEBASEERD OP GELIJKE KLASSEBREEDTE (0,4 PH EENHEID) EN OP GELIJKE AANTAL WAARNEMINGEN (10-PERCENTIELEN)

klasse	Klassen gelijke breedte		Klassen gelijke omvang	
	ondergrens	bovengrens	ondergrens	bovengrens
1	5.4	5.8	5.40	7.20
2	5.8	6.2	7.20	7.45
3	6.2	6.6	7.45	7.60
4	6.6	7.0	7.60	7.72
5	7.0	7.4	7.72	7.83
6	7.4	7.8	7.83	7.93
7	7.8	8.2	7.93	8.05
8	8.2	8.6	8.05	8.20
9	8.6	9.0	8.20	8.41
10	9.0	9.17	8.41	9.10

FIGUUR 3.5

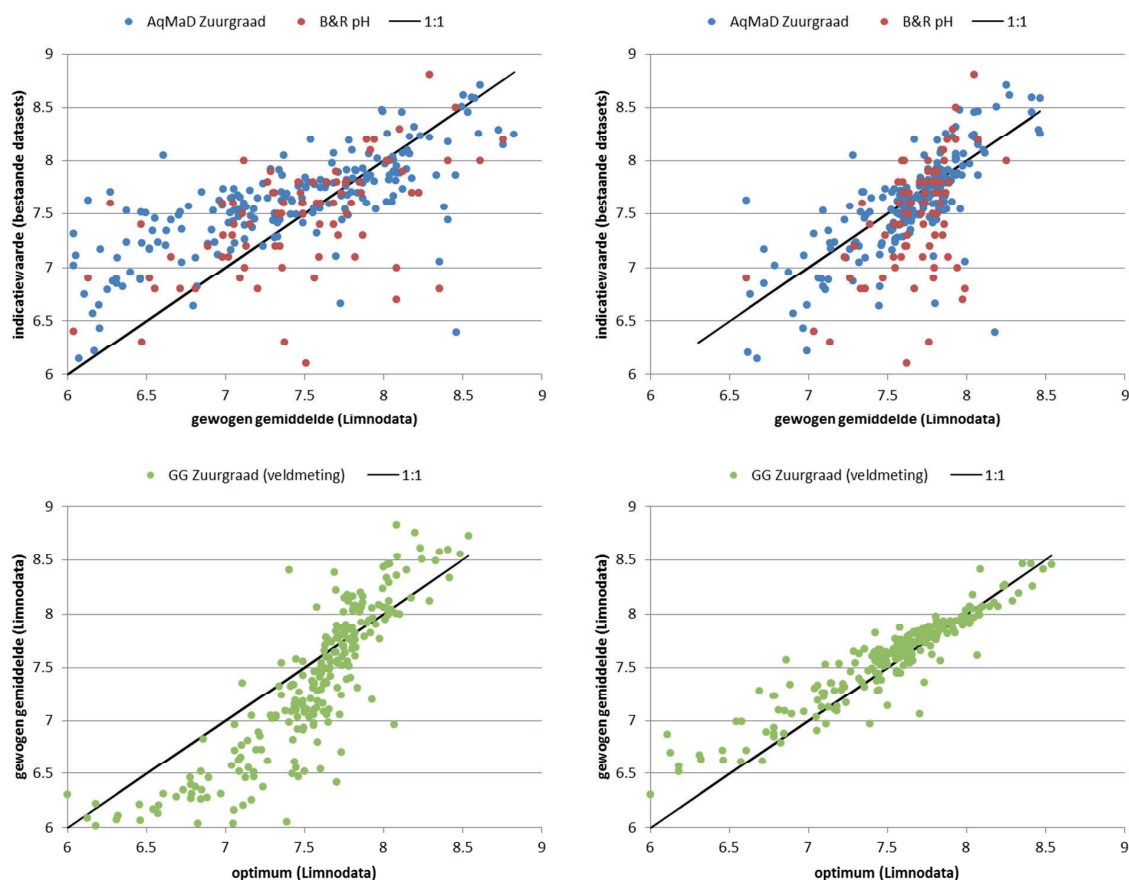
LINKS: KLASSENGRENZEN GG GEBASEERD OP GELIJKE KLASSEBREEDTE VAN DE ZUURGRAAD (0,4 PH EENHEID) EN RECHTS OP GELIJKE AANTAL WAARNEMINGEN PER KLASSE (10-PERCENTIELEN). BIJ EEN GELIJK AANTAL WAARNEMINGEN PER KLASSE VARIEERT DE KLASSEBREEDTE OVER DE HELE RANGE STERK.



Het gewogen gemiddelde (GG) wordt bepaald door 1) het gemiddelde van alle meetwaarden van een milieuvariabele die binnen een klasse vallen te bepalen en deze 2) te wegen met de relatieve frequentie van voorkomen van de soort in die klasse (zie Bloemendaal & Roelofs, 1988). Dus hoe vaker een soort (relatief) voorkomt in een klasse (2), hoe zwaarder het gemiddelde van die klasse (1) weegt.

Figuur 3.6. laat het verschil zien tussen de indeling van klassen op basis van gelijke breedte (links) en van gelijke omvang qua aantal waarnemingen (rechts). Vooral de vergelijking van optimum en gewogen gemiddelde (onderste figuren) laat zien wat er gebeurt: bij klassen van gelijke breedte (linksonder) gaan waarnemingen in klassen met weinig waarnemingen zwaarder meetellen, in dit geval vooral in het lage pH bereik, maar ook in het hoge bereik. Bij klassen met een gelijk aantal waarnemingen (rechtsonder) is er weinig verschil tussen het gewogen gemiddelde en het optimum.

FIGUUR 3.6 VERGELIJKING VAN HET GG VOOR DE ZUURGRAAD (VELDMETING) GEBASEERD OP GELIJKE KLASSEBREEDTE (LINKS, 0.4 PH EENHEID) MET HET GG GEBASEERD OP EEN GELIJK AANTAL WAARNEMINGEN PER KLASSE (RECHTS, 10-PERCENTIELEN). DE BOVENSTE FIGUREN LATEN DE VERGELIJKING VAN HET GEWOGEN GEMIDDELDE ZIEN MET DE INDICATIEWAARDE UIT BESTAANDE DATASETS. DE ONDERSTE FIGUREN DE VERGELIJKING VAN HET GG MET HET OPTIMUM UIT DE LIMNODATA.



Deze bevindingen hebben er toe geleid voornamelijk uit te gaan van de optima (naar abundantie gewogen gemiddelden). De berekening hiervan is eenduidig, waardoor deze de meest robuuste indicatiewaarden opleveren. Wel zijn ze gevoelig voor de verdeling van de data in de dataset. Dit pleit eigenlijk voor het ontwikkelen van een “hybride” methode, die de voordelen van beide methoden benut. Verder is het uiteraard van belang om niet alleen vanuit oogpunt van statistiek naar de data te kijken, maar ook naar relevante ecologische grenzen (“drempelwaarden”) en hier de klassengrenzen op te baseren.

3.5 ALTERNatieve METHODE VOOR BEPALEN MILIEUCONDITIES UIT SOORTEN

Gezien bovenstaande bevindingen is een “alternatieve” methode bedacht en onderzocht om de milieukondities af te leiden uit de voorkomende soorten. Dit is deels afgeleid van de methode Bloemendaal & Roelofs, maar kiest een andere insteek. Overeenkomst is dat de methode eveneens werkt met klassen van een milieuvariabele. Het idee is dat op basis van de data in de trainings-set, de **kans op voorkomen** (relatief getal tussen 0 en 1) wordt bepaald per soort en per klasse van een milieuvariabele. Deze is in principe onafhankelijk van de verdeling van het aantal waarnemingen over de klassen, mits er uiteraard voldoende waarnemingen in iedere klasse zijn. Bij een gering aantal waarnemingen gaat het toeval een rol spelen. Bij de indeling in klassen moet hiermee rekening worden gehouden waarbij tevens wordt aanbevolen te streven naar een evenredige verdeling van klassen over de gradiënt (gelijke breedte) en rekening te houden met ecologisch

relevante grenzen. Hoeveel waarnemingen er minimaal per klasse nodig zijn, is niet in het algemeen te zeggen.

Figuur 3.7. geeft een voorbeeld voor enkele soorten en parameters uit de Limnodata dataset. De figuur laat de kans op voorkomen (KOV) zien per soort en per klasse (in totaal 10 klassen) van de betreffende milieuvariabele. NB! De schaling van de kans op voorkomen is relatief, ten opzichte van de hoogste KOV.

In de meeste gevallen is er een duidelijk patroon waarneembaar, met een negatieve trend, een positieve trend of eerst een toename en daarna een afname van de kans op voorkomen over de gradiënt. Hoe duidelijker het patroon en hoe groter de verschillen in de kans op voorkomen, hoe aannemelijker dat de milieugradiënt ook echt sturend is voor het voorkomen van die soort. NB! Er kan echter altijd sprake zijn van een correlatie of interactie met andere milieuv variabelen.

Omgekeerd kan uit de figuur worden afgeleid dat wanneer in een water gele plomp (*Nuphar lutea*) wordt aangetroffen, de kans het grootst is dat het bicarbonaatgehalte in klasse 7 ligt (in dit geval tussen 72 en 124 mg/l). Echter ook de naastgelegen klassen hebben een relatief hoge kans. De klassen 1 t/m 5 hebben echter een lage kans. Door de kansen op voorkomen van meerdere soorten (in dit voorbeeld *Elodea nuttallii* en *Azolla filiculoides*) samen te “combineren”, kan dit verder worden aangescherpt. Dit is gedaan in de laatste kolom met staafdiagrammen (samen). De schatting voor het bicarbonaatgehalte is dan bijvoorbeeld het midden van de klasse met de hoogste kans, 286 mg/l. De figuur geeft echter ook aan dat er 4 klassen zijn met een relatief hoge kans, de soorten indiceren een waarde die het meest waarschijnlijk ergens in dit bereik ligt (tussen circa 42 en 360 mg/l). Binnen dit bereik komen de soorten vaak voor en is het lastig verder te differentiëren. Het gebruik van een range maakt het enerzijds wat moeilijker toepasbaar (géén vast getal maar een range), maar doet wel meer recht aan het feit dat soorten nou eenmaal binnen een range aan omstandigheden kunnen voorkomen.

Het verschil met de methode van Bloemendaal & Roelofs zit vooral in de wijze waarop de indicatiewaarde wordt bepaald (1) en in de presentatie (2) van het resultaat:

- 1 Bij het bepalen van de indicatiewaarde door Bloemendaal en Roelofs, spelen waarnemingen over de gehele gradiënt een rol. Hiertussen zitten soms ook waarnemingen in klassen met extreme hoge of lage waarden (bijvoorbeeld een waarneming met een totaal-P gehalte van 1 mgP/l telt 10 x zo zwaar mee als een waarneming van 0.1 mgP/l). Deze kunnen het gewogen gemiddelde soms sterk beïnvloeden. Bij de alternatieve methode wordt de indicatiewaarde afgeleid van de klasse(n) met de hoogste kans op voorkomen. In het voorbeeld is dit telkens één klasse (met de hoogste KOV), er kan ook gekozen worden voor het gemiddelde van de klassen met een KOV van minimaal 80% van de maximale KOV;
- 2 Het (per klasse) presenteren van de kans op voorkomen van de soort over de gradiënt is zeer inzichtelijk (zie bijvoorbeeld figuur 3.7). Bloemendaal & Roelofs presenteren dit niet, hoewel het wel uit de data in hun tabellen is af te leiden.

FIGUUR 3.7

VERDELING VAN DE KANS OP VOORKOMEN VAN DRIE SOORTEN OVER DE KLASSEN VAN EEN MILIEUGRADIËNT. IN DE KOLOM "SAMEN" IS DIT GECOMBINEERD VOOR DE DRIE SOORTEN. DE KOLOM "SCHATTING" IS HET MIDDELEN VAN DE KLASSE MET DE HOOGSTE KANS OP VOORKOMEN VAN DE SOORTCOMBINATIE (SAMEN).

parameter	<i>Nuphar lutea</i>	<i>Elodea nuttalli</i>	<i>Azolla filiculoides</i>	samen	schatting (klasse max KOV)
Bicarbonaat					286
Calcium					91
Chloride					51
Doorzicht					0.74
Fosfor-totaal					0.76
Kjeldahl stikstof					1.44
Nitraat					0.54
ortho-Fosfaat					0.25
Stikstof-totaal					2.35
Sulfaat					47
Zuurgraad (veldmeting)					7.81
Zuurstofverzadigingspercentage					59

In figuur 3.8. nog twee voorbeelden van soortcombinaties die een lage pH, bicarbonaat, calcium, chloride en sulfaat indiceren (links) en soorten die een lage totaal P- en N-gehaltes en een hoog doorzicht indiceren (rechts).

De methode (waarbij de MIW wordt afgeleid van de klasse(n) waar het relatieve voorkomen het grootst is) is nog onvoldoende getest om te kijken of het een betere voorspellende waarde geeft dan de voorgaande methoden (ter Braak en Bloemendaal & Roelofs). Hiertoe zou wellicht ook de abundantie van soorten meegewogen moeten worden. Wel is al duidelijk dat het een veel beter inzicht geeft in de spreiding rondom een geschatte waarde van een milieuvariabele. Daarnaast kan de informatie ook gebruikt worden voor het schatten van de kans op voorkomen van een soort, als functie van een milieuvariabele. Daarvoor kunnen overigens ook andere typen responsfuncties (zoals logistische regressie) worden gebruikt. Het schatten van de kans op voorkomen van soorten bij een bepaalde waarde van een milieuvariabele kan bijvoorbeeld worden gebruikt voor de inschatting van het effect van maatregelen op de EKR-score via de soorten.

FIGUUR 3.8

RELATIEVE VERDELING VAN DE KANS OP VOORKOMEN VAN TWEE SOORTCOMBINATIES DIE SPECIFIEKE OMSTANDIGHEDEN INDICEREN (ZIE TEKST).

parameter	Hydrocotyle vulgaris, Juncus bulbosus en Sphagnum cuspidatum	schatting (klasse max KOV)	Chara vulgaris, Elodea canadensis en Hottonia palustris	schatting (klasse max KOV)
Bicarbonaat		6		98
Calcium		3		53
Chloride		16		29
Doorzicht		0.28		0.93
Fosfor-totaal		0.11		0.11
Kjeldahl stikstof		0.57		0.57
Nitraat		0.18		0.18
ortho-Fosfaat		0.08		0.08
Stikstof-totaal		0.96		0.96
Sulfaat		8		30
Zuurgraad (veldmeting)		5.58		7.07
Zuurstofverzadigingspercentage		105		59

4

SYNTHESE

4.1 DISCUSSIE EN CONCLUSIES

De analyse heeft veel informatie opgeleverd over de afleiding en toepassing van milieu-indicatiewaarden. Duidelijk is dat de vele keuzes die moeten worden gemaakt bij het afleiden van de indicatiewaarden en het afleiden van de voorspellingmodellen, grote invloed hebben op de uitkomsten. Van belang zijn onder andere de keuze van de dataset (alle data of indelen in deelsets), keuze van soorten (alleen submers, drijfblad en kroos of alle KRW-taxa), het wel of niet toepassen van transformatie van abundanties van soorten en waarden van milieuvariabelen (LN of wortel of nog iets anders) en de methode (optimum of gewogen gemiddelde volgens Bloemendaal & Roelofs). Verder is van belang hoe de spreiding in waarnemingen is; evenredig over de milieugradiënt of juist geclusterd en in hoeverre wordt de hele milieugradiënt afgedekt? Hierover wordt het volgende opgemerkt (**conclusies / concluderende bevindingen zijn vetgedrukt**):

- *Keuze van de dataset*: de analyses laten zien dat er duidelijke verschillen zijn in respons van soorten (dezelfde soort komt onder andere omstandigheden voor) tussen watertypen, landschapsregio's en waterbeheerders. Tegelijkertijd lijkt het er sterk op dat de patronen wel vergelijkbaar zijn (soorten reageren op een vergelijkbare manier). Tevens laten de analyses zien dat data van verschillende watertypen elkaar aanvullen, voor wat betreft de dekking van de milieugradiënt, en omgekeerd, de data van individuele watertypen of regio's niet de hele gradiënt afdekken. Ook zijn de resultaten bij gebruik van de gehele dataset niet per sé slechter! **Dit pleit dus voor het bepalen van indicatiewaarden en voorspellingmodellen op één of enkele grotere sets (bijvoorbeeld onderscheid op basis van stilstaand/stromend, zoet/brak, laag/hoog NL) in plaats van vele kleine.** Ook vanuit praktisch oogpunt zijn één of enkele grotere sets te preferen boven meerdere kleine.
- *Keuze van de soorten en minimaal aantal waarnemingen*: de verwachting was dat de echte watergebonden soorten (submers, drijfblad en kroos) het beste resultaat zouden laten zien voor wat betreft de voorspelling van milieucondities. Uit de analyse bleek echter dat de totale KRW-lijst beter presteerde. **Dit pleit er voor om alle KRW-soorten mee te nemen.** Wel kan worden overwogen om per soort en milieuvariabele een minimum aantal waarnemingen te stellen, bijvoorbeeld 10. Daartoe is nog een vergelijkende analyse gedaan op de gehele dataset (zie bijlage 6). Hieruit komt het volgende naar voren: **de voorspelling (voor de test-set) wordt in het algemeen slechter wanneer soorten met minder dan 10 waarnemingen worden uitgesloten.** Dit zijn juist vaak indicatieve (maar ook zeldzamere) soorten die aan de randen van de gradiënt worden aangetroffen. **De conclusie is dus dat het beter is deze "zeldzame" soorten wel mee te nemen, echter nog beter is om gericht te kijken of gegevens van dergelijke milieus aan de dataset kunnen worden toegevoegd** (ten behoeve van een meer evenredige verdeling van waarnemingen over de gradiënt).
- *Spreiding in waarnemingen*: dit is een lastig punt en hangt deels samen met de keuze van de dataset (hoe groter de dataset, hoe groter de kans dat de gradiënt gedekt zal zijn) en met de

transformatie (transformatie lost vaak al een deel van de onevenredige spreiding op). Toch verdient het nadere aandacht, omdat bij het bepalen van de optima en het afleiden van de voorspellingsmodellen, de verdeling van de data erg bepalend is voor de uitkomst. Dit leidt tot een slechtere prestatie van de voorspelmodellen. **Dit pleit er voor om te streven naar een dataset met een evenredige verdeling van waarnemingen over de gradiënt.**

- *Transformatie*: het toepassen van transformatie op abundanties van soorten en waarden van milieuvariabelen heeft een positief effect op de resultaten van de voorspelling, met name het voorkómen van negatieve waarden. LN-transformatie levert in de meeste gevallen het beste resultaat. **Dit pleit voor het standaard toepassen van LN(x+1) transformatie**, met uitzondering van pH (want die is reeds getransformeerd). Aandachtspunt is dat het de analyse en bewerking wat complexer maakt, waardoor het risico op fouten ook toeneemt. Ook is een LN-getransformeerde indicatiewaarde minder (niet direct) informatief. Er kan wellicht onderscheid gemaakt wordt tussen de voorspelmodellen waar (achter de schermen) met getransformeerde waarden wordt gewerkt en gepubliceerde indicatiewaarden die ongetransformeerd (of teruggetransformeerd) zijn.
- *Keuze van de methode voor bepalen van indicatiewaarden*: de berekening van het gewogen gemiddelde volgens de methode “Bloemendaal & Roelofs” lijkt een oplossing te zijn voor een ongelijkmatige verdeling van de data over de milieugradiënt, maar is in de praktijk gevoelig voor de keuze van de klassengrenzen. De volgende twee - reproduceerbare - methoden om de klassengrenzen te kiezen zijn onderzocht; klassen met gelijke breedte of klassen met gelijk aantal waarnemingen:
 - Bij klassen met een gelijke breedte (traject van de milieuparameter) komt het vaak voor dat er enkele klassen zijn met een gering aantal waarnemingen, die een grote invloed kunnen hebben op het gewogen gemiddelde. Soms is er één klasse waarin de bulk van het aantal waarnemingen valt. Dit leidt soms tot absurde resultaten;
 - Bij klassen met een gelijk aantal waarnemingen (percentielwaarden) lijkt de uitkomst vaak erg op het optimum, en gaat het sterke punt van deze methode verloren.

Beide uitkomsten zijn ongewenst. Klassengrenzen op basis van ecologisch relevant geachte grenswaarden zijn mogelijk een oplossing. Ook deze zijn toegepast, maar zijn niet voor alle parameters even duidelijk (o.a. wel voor pH, bicarbonaat en calcium maar niet voor doorzicht). Ook kunnen ze verschillen tussen watertypen en/of landschaps-ecologische regio's (bijvoorbeeld voor totaal-P). Al met al zijn ze dus nog niet zo eenvoudig objectief en daarmee reproduceerbaar af te leiden. Hoe dan ook, **de keuze voor de klassengrenzen heeft grote invloed op de waarde van het gewogen gemiddelde met de methode Bloemendaal & Roelofs**, dit maakt een reproduceerbare toepassing van deze methode lastig, **dit pleit er voor om gebruik te maken van de optima**. Het verdient echter aanbeveling om te kijken of er nog een tussenoplossing is, die de sterke punten van beide methoden benut en objectief/reproduceerbaar is. Hiertoe is al een **alternatieve methode bedacht en onderzocht die uitgaat van de kans op voorkomen van een soort per klasse van een milieuvariabele**. Deze methode lijkt kansrijk, maar behoeft en verdient nog verdere uitwerking (zie hieronder).

“Één indicatiewaarde” of alternatief: “kans op voorkomen over de range van een milieugradiënt”

Zoals blijkt uit de analyses zitten er veel haken en ogen aan het gebruik van één enkele indicatiewaarde. Soorten worden vaak over een substantieel deel van een milieugradiënt aangetroffen, waarbij de kans op voorkomen afhankelijk is van de plek op de gradiënt. Door hieruit één indicatiewaarde af te leiden, raakt veel informatie verloren. Door de kans op voorkomen

(per klasse) over de gehele range van de milieuvariabele te gebruiken, blijft deze informatie wél behouden. Tegelijkertijd maakt dit het toepassen van de gevonden relaties van soorten met milieuvariabelen echter wel complexer. Ook doet het geen recht aan de verschillen in abundantie; of een soort met een bedekking van 1% of 90% voorkomt is wel degelijk ecologisch relevant. Dit geldt vooral voor algemenere soorten, zoals waterpest, grof hoornblad of bultkroos. Het wel of niet vóórkomen van deze soorten is vaak maar beperkt indicatief, ze kunnen onder zeer uiteenlopende condities worden aangetroffen. De mate van bedekking is echter zegt vaak veel meer over de milieuomstandigheden.

4.2 AANBEVELINGEN

Deze studie had ten doel om enkele methoden en keuzemogelijkheden bij het afleiden van milieu-indicatiewaarden te vergelijken. Gedurende de rit bleek dat er een aantal haken en ogen zitten aan het afleiden en toepassen van die indicatiewaarden, dit is in de vorige paragraaf ook bediscussieerd. Afhankelijk van de manier waarop (absoluut of relatief) de indicatiewaarden (of responsfuncties) gebruikt gaan worden, is het al dan niet wenselijk / nodig om datasets en methoden verder te verbeteren / te ontwikkelen.

“Relatief toepassen indicatiewaarden”: Voor het relatief toepassen van milieuindicatiewaarden – dus voor de vergelijking van een opname met een referentielijst of met eerdere opnames op dezelfde plek - zoals dat gebeurt in AqMaD - kunnen bestaande methoden en datasets al worden gebruikt. Hiervoor is het ook niet per sé nodig om indicatiewaarden af te leiden per regio of watertype. Wel wordt aanbevolen onderscheid te maken in stilstaand en stromend. Dit betekent dat de bestaande milieuindicatiewaarden (AqMaD en Bloemendaal & Roelofs), maar beter nog de milieuindicatiewaarden (optima) die worden afgeleid voor stilstaande en stromende wateren uit de Limnodata, goed toepasbaar zijn in AqMaD.

“Absoluut toepassen indicatiewaarden”: Voor het absoluut toepassen van milieuindicatiewaarden – dus voor het doen van een voorspelling van de waarde van een milieuvariabele op basis van de voorkomende soorten - zijn bestaande methoden en datasets nog niet geschikt. Er zijn twee belangrijke problemen:

- Ten eerste is gebleken dat de respons van soorten op een milieuvariabele verschilt tussen watertypen en tussen regio's. Dit levert grote afwijkingen op bij de voorspelling en maakt het nodig datasets in te delen in clusters van wateren die vergelijkbaar zijn qua regio en watertype (zie indeling data in deelsets);
- Ten tweede zijn de methoden gevoelig voor de verdeling van data over de dataset. Een alternatief voor de onderzochte methode (voorspelling op basis van optima), is om te werken vanuit de kans op voorkomen van een soort over een milieugradiënt. Deze methode is veel minder gevoelig voor de verdeling van data. Hiermee kan een voorspelling worden gedaan van een milieuvariabele, waarbij zowel één getal als de range wordt gegeven waarbinnen de soorten mogen worden verwacht. Dit is veel inzichtelijker en verdient nadere aandacht, net als andere kansrijke methoden.

Bovenstaande geldt ook voor het toepassen van responsies in habitat(geschiktheids)modellen.

Bovenstaande betekent ook dat het betrouwbaar afleiden van milieucondities (waarden van milieuvariabelen) op basis van soorten in ieder geval voorlopig moet worden losgelaten. Nog los van de genoemde problemen, zijn er andere belemmeringen bij deze toepassing. Vooral de interactie tussen variabelen (bijvoorbeeld is P, N, C of licht limiterend), het bestaan van “drempelwaardes” (zoals bijvoorbeeld te zien bij respons op pH of calcium) en de invloed

van zowel de milieucondities in de bodem als in het water (planten kunnen bijvoorbeeld zowel nutriënten uit de bodem als het water halen), zorgen er naar verwachting voor dat de gevonden relaties in “absolute zin” zwak zijn.

“Getrapte / gecombineerde toepassing van indicatiewaarden en soortresponsies”: Wellicht liggen er goede potenties in een getrapte toepassing van de indicatiewaarden en soort-milieu-relaties. Denkbaar is een toepassing van “grof” naar “fijn”, door bijvoorbeeld:

- eerst een analyse te maken van mogelijke knelpunten, door het relatief toepassen van de indicatiewaarden (AqMaD). Door te kijken naar het verschil in indicatie van de aanwezige soorten en de gewenste (doel / referentie) soorten, kunnen de meest relevante milieuv variabelen worden geïdentificeerd;
- vervolgens voor deze milieuv variabelen verder in te zoomen op de respons van individuele soorten of soortcombinaties over de gehele gradiënt. Dit geeft een beter inzicht in hoeverre de heersende milieucondities daadwerkelijk beperkend kunnen zijn voor het voorkomen van specifieke soorten. Hiervoor kan een eenvoudige grafische weergave, zoals in figuur 3.8, al zeer behulpzaam zijn. Wel wordt in dat geval aanbevolen uit te gaan van de “best passende relaties”, dat wil zeggen rekening houdend met de watertypen en landschappelijke regio (zie indeling in deelsets);
- tot slot nog een multifactoriële benadering toe te passen. Dit kan op verschillende manieren worden gedaan. Eén benadering is het tegelijkertijd beschouwen van meerdere unifactoriële relaties. Bijvoorbeeld door een matrix van soorten en milieuv variabelen, met de kans op voorkomen van iedere soort bij de heersende milieucondities per milieuv variabele. Een andere manier is om te zoeken naar interacties tussen milieuv variabelen door middel van multivariate (regressie)technieken.

Aanbevolen wordt dit verder uit te werken. Uit bovenstaande (2^e bolletje) volgt dat nader onderzoek naar de respons van soorten (kans op voorkomen over de milieu gradiënt, minimum en maximum grenswaarden) in de verschillende deelsets (op basis van landschapstypen of watertypen) aanvullende waarde kan hebben naast alleen de bepaling van gemiddelde (indicatie) waarde. Op de indeling in deelsets wordt op de volgende pagina ingegaan.

Indeling data in deelsets: Het verdient aanbeveling de datasets in te delen in enkele grotere clusters die zijn gebaseerd op regio (bijvoorbeeld: laag NL marien-beïnvloed, laag NL rivier-beïnvloed en hoog NL) en watertype (stagnant lijnvormig, ondiepe meren, diepe meren, langzaam stromend en snelstromend). Binnen deze clusters moet vervolgens worden gestreefd om de gehele range aan milieugradiënten af te dekken, eventueel met aanvullende datasets. Afhankelijk van de gekozen methode is het van belang om te zorgen voor een evenredige verdeling van waarnemingen over de gradiënt (optima) of voldoende waarnemingen per klasse (GG en kans op voorkomen).

Bovenstaande indeling sluit in hoofdlijnen aan op andere indelingen, o.a. voor de KRW-Verkenner:

- Langzaam stromende beken: Hoog-NL
- Snelstromende beken: Hoog-NL
- Ondiepe meren (onderverdeeld in Laag- en Hoog-NL)
- Diepe meren (onderverdeeld in Laag- en Hoog-NL)
- Zwak brakke wateren: Laag-NL
- Brakke wateren: Laag-NL
- Sloten (zoet) (onderverdeeld in Laag- en Hoog-NL)

LITERATUUR

- Amesz, M. & A. Barendregt, 1995a. Milieu-indicatiewaarden van aquatische macrofauna in Noord-Holland. Deel I, Vakgroep Milieukunde, Universiteit Utrecht. Dienst Milieu en Water & Dienst Ruimte en Groen, Provincie Noord-Holland. Utrecht/Haarlem.
- Amesz, M. en A. Barendregt, 1996. Imram: een voorspellingsmodel voor aquatische macrofauna in Noord-Holland, Faculteit Ruimtelijke Wetenschappen, Universiteit Utrecht.
- Barendregt, A. en M. Wassen, 1989. Het hydro-ecologisch model ICHORS (versies 2.0 en 3.0) - de relaties tussen water- en moerasplanten en milieufactoren in Noord-Holland, Vakgroep Milieukunde, Rijksuniversiteit Utrecht.
- Barendregt, A., J.W. Nieuwenhuis en P. de Joode, 1990. Milieu-indicatiewaarden van water- en oeverplanten in Noord-Holland, Interfacultaire Vakgroep Milieukunde, Rijksuniversiteit Utrecht.
- Birks HJB, Line JM, Juggins S, Stevenson AC, ter Braak CJF (1990a) Diatoms and pH reconstruction. *Philos Trans R Soc B, Biol Sci* 327:263-278.
- Birks HJB & GL Simpson, 2013. 'Diatoms and pH reconstruction' (1990) revisited. *Journal of Paleolimnology*, 2013, Page 1.
- Bloemendaal, F.H.J.L. & J.G.M. Roelofs, 1988. Waterplanten en waterkwaliteit. Stichting KNNV en vakgroep Aquatische Oecologie en Biogeologie van de KUN.
- Dam, H. van (2013): Ontwikkeling module diatomeeën voor Volg- en Stuursysteem en KRW-Verkenner. Rapport 906. Herman van Dam, Adviseur Water en Natuur. Amsterdam. 70p. + digitale bijlagen
- De Lyon, M.J.H. & J.G.M Roelofs (1986) Waterplanten in relatie tot waterkwaliteit en bodemgesteldheid. Deel 1 en 2. Technisch Rapport, Katholieke Universiteit Nijmegen, Nijmegen.
- Ertsen Doesjka, Piet Verdonchot, Rick Wortelboer & Bas van der Wal, 2007. Ristori. Modellen voor het voorspellen van de effecten van maatregelen op de aquatische gemeenschappen van sloten en beken. STOWA- rapportnummer 2007-15. STOWA, Utrecht.
- Jaarsma, N.G., R. Gylstra, G. van Ee en G. van Geest, 2015. Pilot soortresponsies waterschap Rivierenland en hoogheemraadschap Hollands Noorderkwartier. H2O-Online / 27 augustus 2015.
- Steenbergen, H.A., 1993. Macrofauna-atlas van Noord-Holland : verspreidingskaarten en responsies op milieufactoren van ongewervelde waterdieren. Provincie Noord-Holland, Haarlem.
- STOWA, 2014. Ecologische sleutelfactoren. Begrip van het watersysteem als basis voor beslissingen. STOWA -rapportnummer 2014-19.
- Ter Braak, C.J.F., 1996. Unimodal models to relate species to environment. DLO-Agricultural Mathematics Group, Wageningen.
- Verdonchot, 2015. Ecologisch raamwerk voor aquatische ecosystemen. Visie op aquatisch ecosysteem functioneren en afgeleide parameters voor modelontwikkeling en water-beheer. STOWA rapport 2015-19. STOWA, Amersfoort.