

Het gewicht van bedekking

Optimalisatie van de berekening van gewogen gemiddelde indicatiewaarden voor vegetatieopnamen

Jos Käfer

Begeleider: J.P.M. Witte

Rapport van een afstudeervak Agro-/Ecohydrologie

Rapport 112

Sectie Waterhuishouding
Nieuwe Kanaal 11, 6709 PA Wageningen
Internet: www.dow.wau.nl/whh

ISSN 0926-230X

1667563

Samenvatting

Om een beeld te krijgen van de toestand van de standplaatsfactoren in een vegetatieopname is het berekenen van (gewogen) gemiddelde indicatiewaarden een gangbare praktijk. Gewogen gemiddelden worden vaak berekend door de bedekking van elke plantensoort op één of andere wijze te transformeren tot een gewicht voor de indicatiewaarde van de soort. Er is echter slechts in beperkte mate onderzocht welke methode daarvoor het beste is. In deze studie wordt gezocht naar een optimale methode voor het berekenen van gemiddelde indicatiewaarden voor vocht en zuurgraad met de indicatiewaarden van Ellenberg en het ecotopensysteem. Daarbij zijn drie bestaande methoden (geen weging, weging met bedekkingspercentages en weging met behulp van de gemodificeerde Braun-Blanquet schaal) met elkaar vergeleken. Bovendien is er door optimalisatie naar nieuwe, betere methoden gezocht. Voor deze optimalisatie is een algemene bedekkingsweegwaardeformule met enkele vrij te variëren parameters ontwikkeld. De optimale methode is dan de bedekkingsweegwaardeformule met dié combinatie van parameters, waarbij de berekende gewogen gemiddelde indicatiewaarden de hoogste correlatie vertonen met gemeten pH of grondwaterstand. Het vergelijken van methoden gebeurt met een permutatietest, gebaseerd op computer-intensieve *resampling* statistiek. De permutatietest, die bij dit onderzoek zeer geschikt is gebleken, is aangepast en gebruikt om verschillen tussen twee Spearman rangcorrelatiecoëfficiënten te testen.

Bij de zuurgraadindicatie blijkt de transformatie van bedekkingen in gewichten niet veel uit te maken. Voor de vochtindicatie met beide indicatiesystemen is een betere dan de bestaande methoden gevonden. In deze betere methode is het verschil tussen laagste en hoogste bedekkingsgewicht klein: in tegenstelling tot veel gebruikelijke methoden, krijgen soorten met de hoogst mogelijke bedekking een gewicht dat slechts twee tot drie keer zo groot is als het gewicht van soorten die met slechts één exemplaar in de opname voorkomen. Bovendien bleken de gemiddelde ecotopensysteem-indicatiewaarden beter overeen te komen met de grondwaterstand dan de gemiddelde Ellenberg vochtgetallen.

Uitgaande van het feit dat de bedekking die een soort bereikt sterk bepaald wordt door de specifieke groeivorm van de soort, is in deze studie geëxperimenteerd met verschillende manieren om het gewicht van elke soort af te laten hangen van de gemiddelde of maximale bedekking van die soort. Dit leverde geen duidelijk betere resultaten op.

Bovendien zijn enkele alternatieve manieren van wegen, die niet direct betrekking hebben op de bedekking van soorten, ontwikkeld en onderzocht. Om het onderscheidingsvermogen van gemiddelde waarden te verhogen, kunnen soorten met erg hoge of lage indicatiewaarden (soorten van extreme milieus) een hoger gewicht krijgen, afhankelijk van de hoeveelheid indifferente soorten in de opname. Bij de vochtindicatie kunnen de resultaten van de Ellenberg indicatiewaarden op deze wijze significant verbeterd worden, bij de zuurindicatie niet. Een andere mogelijkheid, die

vaak geopperd maar slechts zelden gebruikt wordt, is het toekennen van grotere gewichten aan soorten naarmate hun ecologische amplitude nauwer is. De ecologische amplitudes, zoals die met het ecotopensysteem beschreven zijn, leverden voor dit doel geen duidelijke verbetering van de correlatie van indicatiewaarden met het milieu op.

Voorwoord

Dit rapport is het product van een afstudeervak Agro-/Ecohydrologie bij de sectie Waterhuishouding van Wageningen Universiteit, begeleid door Flip Witte. Wat begon als een eerste afstudeervak van 5 maanden 'om de winter door te komen' liep uit op een onderzoek van heel wat meer maanden met vele mogelijkheden en zijpaden die verkend moesten worden. Niet alleen ben ik nu flink wat ervaring en kennis rijker, ook heeft dit afstudeervak mijn kijk op de ecologie en wetenschap veranderd. Ik heb mij niet hoeven vervelen, en het grootste deel van de tijd met plezier achter de PC of met de neus in de boeken gezeten. De voornaamste van de hier beschreven conclusies zullen nog verwerkt worden tot een tijdschriftartikel. Ik hoop dat dit rapport of het artikel dat hieruit voort zal vloeien een nuttige bijdrage levert aan het gebruik van indicatiewaarden, en mogelijk de aanleiding is tot een nadere beschouwing van de hier geschetste problemen en oplossingen.

In de inleiding heb ik, naast de doelstelling, iets van de benodigde voorkennis beschreven, met als gevolg dat enkele onderdelen triviaal over kunnen komen op in dit vakgebied bekende lezers (vooral de beschrijving van de methodiek van de vegetatieopname en de beschrijvingen van het indicatiewaardensysteem van Ellenberg en het ecotopensysteem). De inleiding biedt wel een aanknopingspunt voor de rest van dit rapport (vooral paragraaf 1.4.1). De beschrijving van de methoden in hoofdstuk 2 is tamelijk uitgebreid. Hierin bespreek ik de concepten en meningen van andere auteurs die uiteindelijk hebben geleid tot de hier gepresenteerde methoden. Hoofdstuk 3 met de resultaten bevat ook enkele conclusies en beschouwingen per methode. De discussie in hoofdstuk 4 bevat allereerst een rationale voor het gebruik van *resampling* statistiek. De mogelijke implicaties van de resultaten worden besproken, alsmede de geldigheid van deze studie en enkele relaties met resultaten van andere auteurs. Tot slot bevat de discussie een overzicht van de belangrijkste conclusies. In de bijlagen is onder andere een overzicht van de formules waar vaak naar verwezen wordt te vinden.

Dank ben ik in de eerste plaats verschuldigd aan Flip Witte, die dit onderwerp als afstudeervak aanbood, mij op een (naar mijn mening) prettige wijze begeleid heeft en mij in verschillende ecologische disciplines op een kritische wijze wegwijs heeft gemaakt. Ook wil ik mijn dank betuigen aan Paul Torfs, ook werkzaam bij de sectie Waterhuishouding van Wageningen Universiteit, voor het beantwoorden van mijn vragen op statistisch gebied en het geven van enkele zeer nuttige suggesties. Camiel Aggenbach van Kiwa Water Research hielp bij het uitvoeren van de analyse met het kiwa-indicatorsoortensysteem, en hij en de andere ecologen van Kiwa leverden commentaar en suggesties tijdens mijn aanwezigheid op Kiwa, en tijdens de presentatie van de resultaten. Voor het beschikbaar stellen van gegevens dank ik Han Runhaar (Alterra), André Schaffers (leerstoelgroep Plantenecologie en Natuurbeheer, Wageningen Universiteit) en Stefan Hennekens (Alterra).

Wageningen, oktober 2002

Inhoud

Samenvatting	3
Voorwoord	5
Inhoud	7
Lijst van afkortingen en symbolen	9
1 Inleiding	11
1.1 Achtergronden en doel van deze studie.....	11
1.2 Het maken van een vegetatieopname.....	11
1.3 Indicatiewaarden.....	12
1.3.1 Ellenberg.....	13
1.3.2 ecotopensysteem.....	16
1.4 Gebruik van indicatiewaarden in vegetatieopnamen: bestaande methoden... 17	
1.4.1 Bedekking en gemiddelde indicatiewaarde.....	18
1.4.2 Statistisch correcte methoden.....	19
2 Materialen en methoden	21
2.1 Algemene werkwijze.....	21
2.2 Gebruikte gegevens.....	21
2.2.1 Vegetatieopnamen.....	21
2.2.2 Grondwater.....	22
2.2.3 Zuurgraad.....	23
2.3 Berekening van indicatiewaarden per opname.....	24
2.3.1 Bedekkingsweegwaarden.....	24
2.3.2 Soortspecifieke bedekkingsweegwaarden.....	28
2.3.3 Ellenberg-indicatiewaarden en indicatieweegwaarden.....	30
2.3.4 Ecotopensysteem en soortweegwaarden.....	33
2.3.4.1 Van ecologische groepen naar soorten.....	33
2.3.4.2 Soortweegwaarden.....	35
2.3.5 Mediaan.....	39
2.4 Statistische verwerking.....	40
2.4.1 Correlatie.....	40
2.4.2 Schatting van de correlatie met behulp van de bootstrap.....	41
2.4.3 Het vergelijken van twee correlatiecoëfficiënten met behulp van een permutatietest.....	42
3 Resultaten	45
3.1 Bedekkingsweegwaarden.....	45
3.1.1 Algemeen.....	45
3.1.2 Vochtindicatie.....	45
3.1.2.1 Optimalisatie <i>BWW</i> -functie.....	45
3.1.2.2 De beste methode.....	47
3.1.3 Zuurgraadindicatie.....	49
3.1.3.1 Optimalisatie <i>BWW</i> -functie.....	49
3.1.3.2 De beste methode.....	49

3.2	Ellenberg versus ecotopensysteem	52
3.3	Soortspecifieke bedekkingsweegwaarden	53
3.4	Indicatieweegwaarden (Ellenberg)	55
3.5	Soortweegwaarden (ecotopensysteem)	58
3.6	Mediaan	60
4	Belangrijkste conclusies en discussie	61
4.1	Enkele opmerkingen over statistiek	61
4.2	Het berekenen van gemiddelde indicatiewaarden	62
4.3	Belangrijkste conclusies	66
	Literatuur	69
	Bijlage 1 Overzicht van veel gebruikte formules	73
	Bijlage 2 Gevoeligheid van de <i>BWW</i>-formule voor de parameters.	77
	Bijlage 3 Gevoeligheid van de <i>IWW</i>-formule voor de parameters	85
	Bijlage 4 Standplaatsindicatie met het Kiwa-indicatorsoortensysteem	87
	Bijlage 5 Voorbeeld permutatietest	89

Lijst van afkortingen en symbolen

Uitgezonderd symbolen uit paragraaf 2.4 (Statistische verwerking).

$a_{...}, b_{...}, c_{...}, d$: parameters van verschillende weegwaardefuncties
B	: bedekking ($1/100 \times$ bedekkingspercentage)
BWW	: bedekkingsweegwaarde
cm +mv	: centimeters ten opzichte van het maaiveld (eenheid voor grondwaterstand)
F	: weegfactor van een klasse in het ecotopensysteem
GHG	: gemiddelde hoogste grondwaterstand (cm +mv)
GLG	: gemiddelde laagste grondwaterstand (cm +mv)
GMB	: gemiddeld maximale bedekking (als B)
GVG	: gemiddelde voorjaarsgrondwaterstand (cm +mv)
I	: indicatiewaarde
I_c	: centrale indicatiewaarde
IWW	: indicatieweegwaarde
p	: overschrijdingskans (%)
SSB	: soortspecifieke bedekking (als B)
SWW	: soortweegwaarde
r	: Pearson product-moment correlatiecoëfficiënt, tenzij anders vermeld
r_s	: Spearman rangcorrelatiecoëfficiënt

1 Inleiding

1.1 Achtergronden en doel van deze studie

Voor diverse soorten onderzoek, vooral ten behoeve van natuurbeheer en landschapsecologie, kan het nodig zijn om informatie te verkrijgen over het abiotisch milieu, zonder daarvoor kostbare bodemanalyses uit te moeten voeren. Als er gegevens van het verleden nodig zijn, zijn deze analyses zelfs niet meer mogelijk. Omdat planten een nauwe relatie met de bodem hebben en eenvoudig te bemonsteren zijn (door getrainde onderzoekers) kan de vegetatie gebruikt worden als indicator voor de toestand van de bodem. Ook is er veel bekend over de relatie van plantensoorten met het abiotisch milieu, zowel door kwantitatief onderzoek naar deze relatie, als door de ervaring van veldbiologen en vegetatiekundigen. In indicatiewaarden is deze kennis is samengevat. Met een indicatiewaarde wordt de voorkeur van een soort voor een bepaalde omstandigheden aangeven.

De vegetatie wordt in Europa vaak bemonsterd door vegetatieopnamen te maken. Van de plantensoorten die in de opname aanwezig zijn, kan dan een indicatiewaarde opgezocht worden, en uit de indicatiewaarden van de afzonderlijke plantensoorten kan een indicatiewaarde voor de hele opname worden bepaald. Behalve van gegevens over de aan- en afwezigheid van de soorten, kan er ook gebruikt worden gemaakt van informatie over de hoeveelheid waarmee elke soort in een opname vertegenwoordigd is.

Hoe deze hoeveelheid (in het veld vaak geschat als bedekking) meegenomen moet worden in het berekenen van een indicatiewaarde voor de hele opname is slechts matig onderzocht. Empirisch onderzoek waarin nagegaan wordt met welke methode de duidelijkste indicatie verkregen wordt lijkt veelal te ontbreken. In de sporadisch uitgevoerde onderzoeken blijft het bij het vergelijken van enkele (meestal slechts twee) methoden (Schaffers & Sýkora 2000, Böcker et al. 1983, Durwen 1982).

Het doel van dit onderzoek is de optimale methode te vinden waarmee bedekkingen gebruikt kunnen worden bij het berekenen van een indicatiewaarde voor een opname. Hiertoe wordt van zeer veel methoden nagegaan welke de beste, met gemeten milieufactoren overeen komende, resultaten oplevert. Ook komen een vergelijking tussen het indicatiewaarde-systeem van Ellenberg en het ecotopensysteem, en enkele aanvullende mogelijkheden binnen deze systemen aan bod.

1.2 Het maken van een vegetatieopname

In Europa worden vegetatieopnamen meestal volgens een vast stramien gemaakt. Hier wordt slechts een korte schets van de standaardwerkwijze gegeven, meer hierover is te vinden in bijvoorbeeld Dierschke (1994), Schamineé et al. (1995), Shimwell (1971) en Mueller-Dombois & Ellenberg (1974).

Het werk begint met het kiezen van een proefvlak, zodanig dat de vegetatie binnen dit proefvlak homogeen is en een representatief beeld geeft van het te beschrijven plantendek. Dit zijn geen duidelijk meetbare criteria, dus het plaatsen van een proefvlak gebeurt op basis van het oordeel van de (deskundige) onderzoeker. Wel is het mogelijk om enigszins objectief vast te stellen hoe groot het proefvlak moet zijn (zie bovengenoemde publicaties), maar omdat dit een tijdrovend karwei is wordt dit bijna altijd achterwege gelaten. Alle soorten in het proefvlak worden genoteerd, waarna voor elke soort de bedekking geschat wordt. Voor dit schatten van de bedekking kunnen verschillende schalen gebruikt worden, de ene schaal iets verfijnder dan de andere, maar allemaal overeenkomend wat werkwijze betreft. Als voorbeeld volgt hier een beschrijving van de werkwijze met de veel gebruikte gemodificeerde Braun-Blanquet schaal (zie Tabel 4, pag. 25).

Bedekt een soort meer dan 5% van het oppervlak van de opname (loodrechte projectie), dan wordt deze bedekking geschat. Wanneer deze minder dan 5% is, wordt de abundantie (= aantal individuen) geschat. De criteria voor toedeling aan een schattingsklasse en de codering hiervan zijn te vinden in Tabel 4 (op pag. 25, eerste drie kolommen). Het is echter lang niet altijd gemakkelijk om een goede schatting te geven, bijvoorbeeld bij planten met een sprietachtige groeiwijze die verspreid tussen de andere planten in het proefvlak groeien.

Eventueel kunnen nog andere zaken genoteerd worden om de toestand van de vegetatie te verduidelijken, zoals de fenologie, vitaliteit en sociabiliteit, maar dit wordt vaak achterwege gelaten. Ook in dit onderzoek wordt alleen gebruik gemaakt van bedekkingsgegevens.

1.3 Indicatiewaarden

Doordat planten niet allemaal hetzelfde reageren op een milieufactor en meestal slechts voorkomen in een gedeelte van het milieu waar voor de soort geschikte omstandigheden heersen, zijn ze te gebruiken als indicator voor milieuomstandigheden. Het voorkomen van soorten afhankelijk van een milieufactor of een complex van factoren kan dan door middel van een indicatiewaarde aangegeven worden. Deze waarde weerspiegelt dan het voorkomen van een soort ten opzichte van andere soorten (soort a groeit op drogere plekken dan soort b) of kan een directe maat zijn voor de milieufactor (soort a komt voor tussen pH 6 en 7). Er zijn indicatiewaardesystemen waarin alleen maar het ecologisch optimum van een soort aangeven wordt, en systemen die ook informatie verschaffen over de tolerantie van een soort voor verschillende waarden van een milieufactor. Onderzoekers hebben indicatiewaarden opgesteld voor uiteenlopende factoren, variërend van eiwitgehalte van het gras tot hemerobiegraad (Dierschke 1994). Eén van de meest gebruikte lijsten van indicatiewaarden is die van Ellenberg (vanaf 1974).

Het geven van indicatiewaarden aan soorten is een classificatie-methode. Als we er vanuit gaan dat de verspreiding van soortsoptima langs een gradiënt van een milieufactor min of meer continu is (zie bijvoorbeeld Auerbach & Shmida 1993),

zullen de klassengrenzen van elke indicatiewaarde arbitrair vastgesteld moeten worden. Soorten met een overeenkomstige milieuvoorkeur krijgen dezelfde indicatiewaarde en vormen zo samen een ecologische groep.

Punten van kritiek op het gebruik van indicatiewaarden is dat het getalsmatige karakter ervan onjuist gebruik door leken vergemakkelijkt (zie onder andere Durwen 1983), en dat de waarden vaak niet getoetst zijn aan onafhankelijke meetgegevens, maar door deskundigenoordeel tot stand gekomen zijn. Dit laatste kan als een voordeel gezien worden, omdat het voor sommige factoren heel lastig is om de belangrijkste verklarende maat te vinden (Witte & Runhaar 2000). Van de zuurgraad in de bodem is bekend dat deze invloed op de plant heeft via verschillende mechanismen. De pH moet wel erg laag worden voor de H^+ -ionen daadwerkelijk toxisch zijn voor de plant. Bij lage pH is het de verhoogde mobiliteit van aluminium en verminderde beschikbaarheid van voedingsstoffen die schade aan de planten aanrichten. Bij een hoge pH worden ook allerlei voedingsstoffen onbruikbaar voor planten. Beschikbaarheid van voedingsstoffen en mobiliteit van ionen worden echter niet alleen door de zuurgraad bepaald. Het vochtgehalte van de bodem kent ook effecten die afhankelijk zijn van het vochttraject waarin de bodem zich bevindt. Zeer hoge vochtgehalten verminderen de hoeveelheid zuurstof in de bodem, terwijl schade door vochttekorten pas optreedt bij lage vochtgehalten. Dergelijke ingewikkelde relaties kunnen toch vrij goed door onderzoekers en veldbiologen 'aangevoeld' en in indicatiewaarden omgezet worden, wat blijkt uit het succes van bijvoorbeeld het systeem van Ellenberg. Ellenberg (1992) meent, dat het vochtgehalte van de bodem aan grote veranderingen op korte en lange termijn onderhevig is, zodat het onmogelijk is het vochtgetal op basis van meetgegevens te bepalen. Alleen een relatieve indeling is dan mogelijk.

In dit onderzoek worden de indicatiewaarden voor bodemvocht en zuurgraad gebruikt. Het toetsen van de indicatiewaarden is hier geen doel op zich. De indicatiewaarden voor bodemvocht en zuurgraad worden veel gebruikt en de te indiceren factoren zijn redelijk duidelijk. Dat wil echter niet zeggen dat deze factoren ook eenvoudig te meten zijn. Toch zijn de hier gebruikte grondwaterstanden voor vocht en de pH voor de zuurgraad redelijke maten (zie ook paragraaf 2.2). Voor bijvoorbeeld het voedselrijkdom-getal van Ellenberg is veel minder duidelijk wat de belangrijkste verklarende milieufactor is (hierover meer in onder andere Ellenberg (1992), Dierschke (1994), Schaffers & Sýkora (2000)).

1.3.1 Ellenberg

Het systeem van Ellenberg wordt veelvuldig gebruikt. De voor- en nadelen van dit systeem zijn in de literatuur breed uitgemeten (zie o.a. Dierschke 1994, Durwen 1982). Van bijna elke Midden-Europese plantensoort is het optimum geschat en weergegeven door een getal. In de uitgave van 1992 zijn indicatiewaarden voor licht, temperatuur, continentaliteit, vocht (gegeven door het *Feuchtezahl*, F-getal), zuurgraad (bodemreactie, gegeven door het *Reaktionszahl*, R-getal), stikstof (voedselrijkdom), zoutgehalte en resistentie voor zware metalen gegeven, veelal op

een ordinale schaal van 1 tot 9 (maar bij vocht van 1 tot 12). Dit zorgt voor een compleetheid die zeker aan het succes van de lijst bijgedragen heeft. Soorten kunnen maximaal in één klasse ingedeeld zijn. Soorten met een erg brede ecologische amplitude of een ongelijke respons in verschillende geografische gebieden zijn aangegeven als indifferent (aangegeven met 'X').

Bij de indicatie voor vocht is naast de indicatiewaarde van elke soort op een schaal van 1 tot 12 ook weergegeven of een soort bij grote watersstandsschommelingen voorkomt: de '*Überschwemmungszeiger*' komen voor op regelmatig overstroomde grond en '*Wechselzeiger*' indiceren grote schommelingen in de grondwaterstanden. Voor een beschrijving van de verschillende klassen, zie Tabel 1 en Tabel 2.

Omdat de lijst is opgesteld voor het westen van Midden-Europa, is het gebruik van de indicatiewaarden in Nederland enigszins dubieus. Er bestaan mogelijkheden om de indicatiewaarden aan te passen aan het gebied waar ze gebruikt worden (zie o.a. Dierschke 1994, Schaffers & Sykora 2000, Hill et al. 2000). Omdat volledige lijsten met gecorrigeerde indicatiewaarden voor Nederland nog niet algemeen gebruikt worden, is in dit onderzoek uitgegaan van de ongecorrigeerde lijst uit Ellenberg (1992).

De indeling van Ellenberg, vooral de eerste publicatie uit 1974, is gebaseerd op de ervaring van Ellenberg en zijn medewerkers, en in mindere mate op concrete onderzoeken naar de relaties tussen soorten en hun milieu. In de latere publicaties zijn steeds correcties aangebracht, zodat het systeem steeds beter de optima van de soorten weergeeft.

Tabel 1: De vocht-indicatiewaarden uit Ellenberg (1992)

F = Feuchtezahl

Vorkommen im Gefälle der Bodenfeuchtigkeit vom flachgründig-trockenen Felshang bis zum Sumpfboden sowie vom seichten bis zum tiefen Wasser (Nach eigenen Beobachtungen und Angaben von OBERDORFER et al. 1990).

- 1 *Starktrocken*zeiger,
an oftmals austrocknenden Stellen lebensfähig und auf trockene Böden beschränkt
- 2 zwischen 1 und 3 stehend
- 3 *Trocken*zeiger, auf trockenen Böden häufiger vorkommend als op frischen;
auf feuchten Böden fehlend
- 4 zwischen 3 und 5 stehend
- 5 *Frische*zeiger, Schwergewicht auf mittelfeuchten Böden,
auf nassen sowie auf öfter austrocknenden Böden fehlend
- 6 zwischen 5 und 7 stehend
- 7 *Feuchte*zeiger, Schwergewicht auf gut durchfeuchteten, aber nicht nassen Böden
- 8 zwischen 7 und 9 stehend
- 9 *Nässe*zeiger, Schwergewicht auf oft durchnässen (luftarmen) Böden
- 10 *Wechselwasser*zeiger,
Wasserpflanze, die längere Zeiten ohne Wasserbedeckung des Bodens erträgt
- 11 *Wasserpflanze*, die unter Wasser wurzelt,
aber zumindest zeitweilig mit Blättern über dessen Oberfläche aufragt,
oder Schwimmpflanze, die an der Wasseroberfläche flottiert
- 12 *Unterwasser*pflanze, ständig oder fast dauernd untergetaucht
- ~ Zeiger für starken *Wechsel* (z. B. 3~: Wechseltrockenheit,
7~: Wechselseuchte oder 9~: Wechselnässe zeigend)
- = *Überschwemmungs*zeiger,
auf mehr oder minder regelmatig überschwemnten Böden

Tabel 2: De zuurgraad-indicatiewaarden uit Ellenberg (1992)

R = Reaktionszahl

Vorkommen im Gefälle der Bodenreaktion und des Kalkgehaltes (Nach zahlreichen eigenen Messungen und der umfangreichen Literatur sowie nach den Punktrasterkarten im Atlas von HAEUPLER u. SCHÖNFELDER 1989).

- 1 *Starksäure*zeiger, niemals auf schwachsauren bis alkalischen Böden vorkommend
- 2 zwischen 1 and 3 stehend
- 3 *Säure*zeiger,
Schwergewicht auf sauren Böden, ausnahmsweise bis in den neutralen Bereich
- 4 zwischen 3 and 5 stehend
- 5 *Mäßigsäure*zeiger, auf stark sauren wie auf neutralen bis alkalischen Böden selten
- 6 zwischen 5 and 7 stehend
- 7 *Schwachsäure-* bis *Schwachbasen*zeiger, niemals auf stark sauren Böden
- 8 zwischen 7 and 9 stehend, d. h. meist auf Kalk weisend
- 9 *Basen-* and *Kalk*zeiger, stets auf kalkreichen Böden

Wamelink et al. (2002) vonden dat voor elk vegetatietype de relatie tussen gemiddelde Ellenberg-indicatiewaarde en gemeten pH en GVG gegeven werd door een andere lineaire regressielijn, en concludeerden dat gemiddelde Ellenberg-getallen alleen binnen een plantensociologische eenheid vergeleken mogen worden. Zij gingen

echter uit van een lineaire relatie tussen indicatiewaarde en milieufactor. Ellenberg (1992 pag. 44) stelt echter dat de indeling van de soorten niet gebaseerd is op gelijke intervallen tussen twee opeenvolgende indicatiewaarden. In deze studie is de relatie tussen indicatiewaarde en milieufactoren uitgedrukt in de rangcorrelatiecoëfficiënt; deze veronderstelt geen lineair verband. Waarschijnlijk verbetert de rangcorrelatie niet als elk vegetatietype apart geanalyseerd zou worden. Ook Witte & Von Asmuth (in prep.) geven aan dat de conclusie van Wamelink et al. niet getrokken mag worden, omdat deze het gevolg is van het verkeerde statistische model.

1.3.2 ecotopensysteem

Het ecotopensysteem (Runhaar et al. 1987, Stevers et al. 1987, Witte 2002) van het Centrum voor Milieukunde te Leiden (CML) is, in tegenstelling tot het systeem van Ellenberg, toegespitst op de Nederlandse situatie. In het ecotopensysteem zijn bijna alle Nederlandse plantensoorten ingedeeld in ecologische soortengroepen. Deze soortengroepen bestaan uit soorten die karakteristiek zijn voor een bepaald ecotoop. Ecotooptypen zijn ecosysteemtypen, omschreven op basis van milieufactoren en structuur van de vegetatie: zo zijn er de ecotopen 'graslanden op vochtige, voedselarme, zwak zure bodem' (G42) of 'pioniervegetaties op brakke droge stuivende bodem' (bP60st). Deze ecotopen zijn gedefinieerd op basis van structuur, voorkeur voor vochttoestand, zuurgraad, voedselrijkdom, substraat, dynamiek, zoutgehalte en saprobietoestand. Alleen combinaties die men zinnig achtte om de Nederlandse situatie te beschrijven zijn gebruikt om de ecotopen afzonderlijk te definiëren. Zo is binnen de voedselrijke milieus geen onderscheid meer gemaakt naar zuurgraad. De milieukenmerken die gebruikt zijn om elk ecotoop te definiëren zijn in de code van het ecotoop terug te vinden (Tabel 3).

Een belangrijk verschil met het systeem van Ellenberg is dat soorten in meer dan één groep ingedeeld kunnen zijn, waardoor de ecologische amplitude van de soorten enigszins beschreven wordt. Het criterium voor indeling van een soort in één of meer soortengroepen is: "een soort wordt ingedeeld in zoveel soortengroepen als noodzakelijk om minimaal twee derde van het voorkomen van een soort te verklaren, gesteld dat alle ecotooptypen in dezelfde mate zouden voorkomen" (Stevens et al. 1987 pag. 32, Runhaar 1987 pag. 287). De indeling is meer gericht op de standplaats van het hele ecotoop dan het micromilieu van de soort (Stevens et al. 1987). Er is getracht om elke klasse een duidelijke, toetsbare definitie te geven, zodat de indeling minder subjectief is dan die van Ellenberg.

Zoals opgemerkt is, zijn niet alle combinaties zinnig om de Nederlandse situatie te beschrijven. Zo is binnen de brakke en zilte klassen geen onderscheid gemaakt naar voedselrijkdom en zuurgraad, omdat de invloed van het zout zo sterk is dat verschillen in voedselgehalte en zuurgraad geen rol meer spelen. Het tweede cijfer, dat een combinatie van voedselrijkdom en zuurgraad aangeeft, geeft bij matig tot zeer voedselrijke milieus (getallen 7, 8 en 9) geen informatie over de zuurgraad. Dit laatste impliceert dat alle ecologische groepen in voedselrijke milieus als indifferent voor zuurgraad geïnclassificeerd zijn.

Tabel 3: De in het ecotopensysteem onderscheiden kenmerken en kenmerkclassen (naar Runhaar et al. 1987).

saliniteit (prefix)

- zoet

b brak

z zilt

vegetatiestructuur en successiestadium (hoofdletter)

G grasland

H bos en struweel

P pioniervegetatie

R ruigte

V verlandingsvegetatie

W watervegetatie

vochttoestand (eerste cijfer)

1 aquatisch

2 nat

4 vochtig

6 droog

trofietoestand en zuurgraad (tweede cijfer)

1 voedselarm zuur

2 voedselarm zwak zuur

3 voedselarm basisch

7 matig voedselrijk

8 zeer voedselrijk

9 matig tot zeer voedselrijk

0 niet van toepassing

additionele kenmerken (suffix)

kr kalkrijk - alleen binnen P47 en G47

mu stenig substraat - alleen binnen P40 en P60

ro geroerd - alleen binnen P63

sa polysaproob - alleen binnen V18 en W18

st stuwend - alleen binnen bP60

tr betreden - alleen binnen P48

1.4 Gebruik van indicatiewaarden in vegetatieopnamen: bestaande methoden

Voor veel doelen is men niet zozeer geïnteresseerd in de indicatiewaarden van de afzonderlijke plantensoorten, maar in de toestand van het milieu van de gehele vegetatie. Op deze manier kunnen bijvoorbeeld veranderingen van het milieu in de tijd afgeleid worden. Het voorkomen van één enkele plantensoort kan het gevolg zijn van ruimtelijke variatie (variatie op microschaal), temporele veranderingen (bijvoorbeeld naijlende soorten) of toeval. Door gebruik te maken van de informatie uit de gehele vegetatie wordt dan een betrouwbaarder beeld verkregen. Sommigen beweren, dat het milieu van een vegetatieopname, die qua soortensamenstelling

homogeen moet zijn, homogeen is, en dat daarom het milieu ter plekke gelijk moet zijn aan het gebied dat door de maximale overlap van de ecologische amplitudes van alle soorten beschreven wordt (onder andere Schaminée et al. 1995 pag. 226).

De meest eenvoudige en meest gebruikte manier om de indicatiewaarden van de alle soorten in een opname samen te vatten is door middel van de gemiddelde indicatiewaarde van de opname, hoewel andere methoden soms beter geacht worden (mediaan, *Zeigerwertspektrum*).

1.4.1 Bedekking en gemiddelde indicatiewaarde

Naarmate het milieu ter plekke beter met het ecologisch optimum van een soort overeen komt, wordt de concurrentiekracht van een soort groter en kan deze soort hogere bedekkingen bereiken. Indicatiewaarden weerspiegelen in hoge mate dat ecologisch optimum; in het systeem van Ellenberg wordt slechts één indicatiewaarde gegeven die overeen komt met het ecologisch optimum, in het ecotopensysteem wordt het gebied waarin het optimum ligt beschreven door één of meerdere indicatiewaarden (er wordt immers gesteld dat de indeling tweederde van het voorkomen van elke soort in Nederland beschrijft). Bij verreweg de meeste vegetatieopnamen wordt de hoeveelheid van elke soort geschat op een gecombineerde bedekkings- en abundantieschaal. Het ligt voor de hand deze maat, die hier verder simpelweg aangeduid wordt met bedekking, te gebruiken bij het berekenen van een gewogen gemiddelde indicatiewaarde. De algemene formule voor het berekenen van een gewogen gemiddelde indicatiewaarde luidt dan:

$$I_{\text{gem}} = \frac{\sum_{i=1}^N I_i \times BWW_i}{\sum_{i=1}^N BWW_i} \quad (1)$$

waarin

I_{gem} = gewogen gemiddelde indicatiewaarde;

i = de i^{de} soort in de opname;

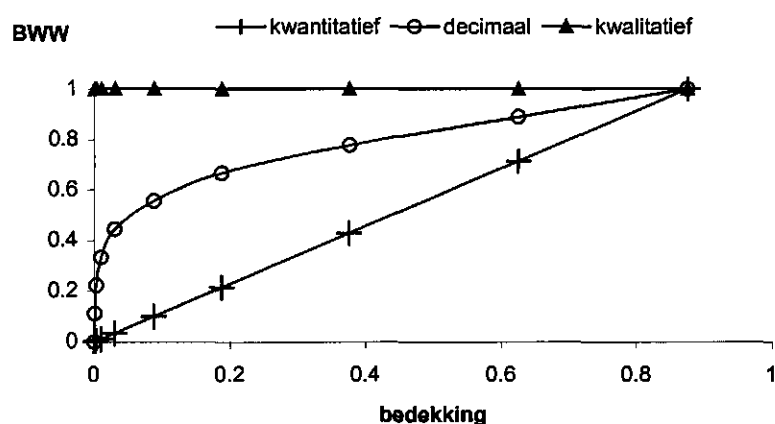
N = aantal soorten in de opname;

I_i = indicatiewaarde van soort i ;

BWW_i = bedekkingsweegwaarde van soort i gebaseerd op de bedekking van soort i .

Om de bedekking van de soorten niet mee te wegen wordt voor alle bedekkingsweegwaarden 1 ($BWW_i = 1$) ingevuld in bovenstaande formule. Ellenberg noemde deze berekening waarbij alleen gebruik wordt gemaakt van de presentie van een soort de kwalitatieve methode (rekenkundig gemiddelde). De eenvoudigste manier om soorten met een hogere bedekking een groter gewicht te geven is door het werkelijke bedekkingspercentage te nemen als bedekkingsweegwaarde (kwantitatieve

methode, $BWW_i = B_i$). Ook is er een lans te breken voor het directe gebruik van de schattingschaal van Braun-Blanquet (Tabel 5) als gewicht. Kleine verschillen in bedekking kunnen immers veelzeggend zijn als de bedekking van de soort in de opname klein is. Bij een hoge bedekking zijn kleine verschillen niet van belang. Omdat vaak een 10-delige schaal gebruikt wordt om de bedekking te schatten (een 9-delige schaal voor soorten die in de opname voorkomen, plus de schatting '0' voor soorten die ontbreken) zal deze methode aangeduid worden met de term decimale methode. De omzetting van bedekkingen naar gewichten volgens de drie hier besproken methoden is grafisch weergegeven in Figuur 1.



Figuur 1: Drie veel gebruikte weegmethoden

1.4.2 Statistisch correcte methoden

Het berekenen van een gemiddelde indicatiewaarde van een vegetatieopname is statistisch niet correct (Durwen 1983, Böcker et al. 1983, Möller 1987, Kowarik & Seidling 1989, Ellenberg et al. 1992, Witte & Runhaar 2000). Indicatiewaarden zijn geen kardinale waarden: een vochtindicatie van 4 is niet twee maal zo droog of nat als vochtindicatie 2. Dit geldt zowel bij het systeem van Ellenberg als bij het ecotopensysteem. Hoewel de mediaan in dergelijke gevallen de correcte centrummaat is, geven de meeste auteurs de voorkeur aan het statistisch niet toelaatbare gemiddelde, omdat de mediaan te veel andere nadelen heeft. De invloed van het toeval is vrij groot op de berekening van de mediaan, want kleine verschillen kunnen er toe leiden dat de mediaan één volledige klasse hoger of lager komt te liggen. Kowarik & Seidling (1989) wijzen in dit verband op een berekeningsmethode voor de mediaan die is beschreven in Sachs (o.a. 1984 pag. 75, zie ook Lienert 1986 pag. 45). De exacte berekeningswijze is te vinden in paragraaf 2.3.5. Voor zover bekend is deze berekening nooit toegepast in een concrete studie waarin indicatiewaarden per opname gebruikt zijn.

Naast de mediaan is de modaal ook een statistisch correcte maat voor het berekenen van een indicatiewaarde van een opname. De modaal is de indicatiewaarde die het

meest voorkomt binnen de opname. Aan deze methode kleven dezelfde nadelen als aan de berekening van de mediaan (Durwen 1982, Kowarik & Seidling 1989).

Het berekenen van een indicatiewaarde voor de hele opname, al dan niet in combinatie met een geschikte maat voor de spreiding van de indicatiewaarden binnen de opname (Durwen 1982), heeft informatieverlies tot gevolg. In een heterogene opname kan de verdeling van de soorten over de indicatiewaarden zodanig zijn, dat het gemiddelde of de mediaan nauwelijks informatie verschaft over het milieu (Durwen 1982). In dat geval kan het nuttig zijn om een staafdiagram van het aantal soorten per indicatiewaarde (of de totale bedekking van de soorten met dezelfde indicatiewaarde) te maken. In de Duitstalige literatuur wordt een dergelijk diagram een '*Zeigerwertspektrum*' genoemd (Ellenberg 1992, Durwen 1982, Kowarik & Seidling 1989). Voordeel hiervan is dat zaken als de spreiding van de indicatiewaarden, eventuele scheve verdelingen, uitbijters e.d. zichtbaar worden. Dergelijke diagrammen kunnen nuttig zijn bij het vergelijken van enkele opnamen. Wanneer echter grotere datasets gebruikt worden, is deze methode onoverzichtelijk en zal de informatie verder gereduceerd moeten worden (bijvoorbeeld tot één getal per opname).

Soorten met dezelfde of overeenkomstige indicatiewaarde kunnen gerangschikt worden in een ecologische groep. Het is dan mogelijk om het aandeel van soorten uit een ecologische groep ten opzichte van het totaal uit te rekenen. Deze methode heeft de formele bezwaren die kleven aan het berekenen van het gemiddelde niet. Voorbeelden van deze berekening worden beschreven in Jansen et al. (2000), Witte & Runhaar (2000), Runhaar et al. (1997) en Möller (1987). De soortengroep waarvan het aandeel berekend wordt zal dan zo gekozen moeten worden dat de uitkomst de variatie in de hele dataset weer kan geven. Als er zowel hele droge als hele natte opnamen aanwezig zijn, kan het aandeel natte soorten niet gebruikt worden om onderscheid te maken tussen de droge en minder droge opnamen (waar geen natte soorten in voorkomen), en het aandeel droge soorten laat geen verschillen zien tussen de natte en minder natte opnamen (zie ook Runhaar & Jansen 1999, Runhaar et al. 1997).

2 Materialen en methoden

2.1 Algemene werkwijze

Voor dit onderzoek stonden vegetatieopnamen met voor elke opname gemeten milieufactoren (paragraaf 2.2) ter beschikking. Om te bepalen hoe goed de verschillen tussen de opnamen, berekend op verschillende wijzen (paragraaf 2.3), overeenkomen met de daadwerkelijke gemeten verschillen in het milieu, wordt steeds de correlatie tussen indicatiewaarden van de opnamen en de gemeten factoren van de opnamen berekend (paragraaf 2.4).

Het grootste gedeelte van het onderzoek is uitgevoerd met programma's die geschreven zijn in Visual Fortran (Fortran 95) met behulp van Microsoft Developer Studio 97, waarbij gebruik is gemaakt van subroutines die eerder geschreven zijn door B. Bleij en F. Witte. Voor enkele aanvullende toetsen en berekeningen zijn het statistisch pakket SPSS en het programma Microsoft Excel gebruikt.

2.2 Gebruikte gegevens

De gegevens zijn niet specifiek verzameld voor dit onderzoek. Wel zijn er enige overeenkomsten in de oogmerken van de onderzoeken waarvoor de gegevens verzameld zijn en dit onderzoek. Er zijn twee bestanden met vegetatieopnamen en metingen gebruikt. André Schaffers stelde 74 opnamen met metingen aan bodem, grondwater en vegetatie ter beschikking die onder meer gebruikt zijn om de betrouwbaarheid van de Ellenberg indicatiewaarden voor vocht, zuurgraad en voedselrijkdom te onderzoeken (door Schaffers & Sýkora (2000)). Van Han Runhaar kwamen 202 opnamen, verzameld in het kader van het project 'Toetsingsonderzoek ecotopensysteem' (Runhaar 1989a, 1989b).

2.2.1 Vegetatieopnamen

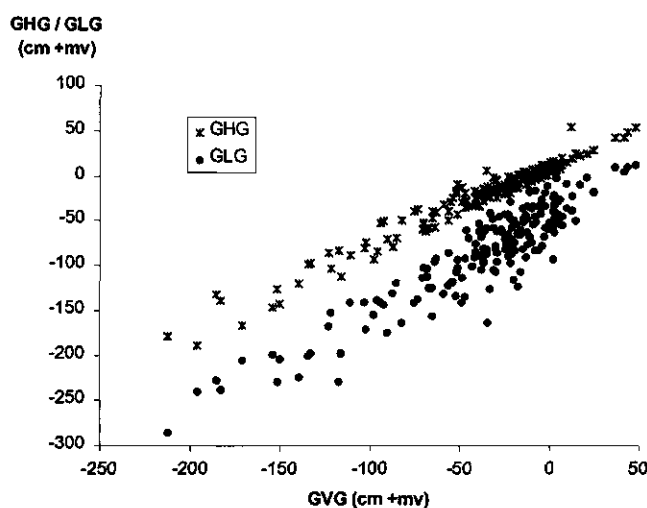
André Schaffers heeft in de zomer van 1992 74 vegetatieopnamen gemaakt voor zijn promotieonderzoek naar de ecologie van wegbermgemeenschappen (Schaffers 2000). De beschrijving is te vinden in Schaffers & Sýkora (2000). De 74 opnamen (verspreid over heel Nederland) zijn geclassificeerd in 14 plantengemeenschappen (van droge kalkgraslanden tot overstromingsgemeenschappen) en omvatten veel variatie in milieuomstandigheden. Van elke gemeenschap zijn meestal 5 opnamen van 25 m² gemaakt. Alleen mossen en vaatplanten zijn genoteerd, de aangeplante bomen en struiken niet.

Een uitgebreide beschrijving van de onderzoeksgebieden waarin Han Runhaar opnamen gemaakt heeft is te vinden in Runhaar (1989b). In de gebieden hebben geen ingrijpende veranderingen in de waterhuishouding plaatsgevonden, en ze zijn zo geselecteerd dat ze een zo groot mogelijke verscheidenheid aan bodemtypen en

vegetatietypen vertegenwoordigen. Gestreefd is naar een proefvlakgrootte van 25 m². De kleinste opname besloeg 2 m² en de grootste 400 m².

2.2.2 Grondwater

Een aantal logische maten voor de grondwaterstand staan ter beschikking voor het berekenen van correlaties met de indicatie van de vegetatie. Volgens Runhaar (1989a) en Runhaar et al. (1997) is de gemiddelde voorjaarsgrondwaterstand (GVG) het meest bepalend voor de plantengroei. Alkemade et al. (1996) maakten ook gebruik van de gegevens van Runhaar (1989b) en vonden de hoogste verklaarde variatie (r^2) bij de regressie van gemiddelde Ellenberg indicatiewaarde met de GVG. Schaffers & Sýkora (2000) vonden dat de gemiddelde grondwaterstand, de gemiddeld hoogste grondwaterstand (GHG) en de GVG een ongeveer even goede correlatie met de indicatie van de vegetatie hebben. In de gegevens van Runhaar blijken GVG en GHG in hoge mate gecorreleerd te zijn (Spearman rangcorrelatiecoëfficiënt (r_s) = 0,964). De correlaties van GHG en GVG met GLG zijn minder hoog (r_s = 0,778 resp. r_s = 0,857; zie ook Figuur 2). Om de weegmethoden te testen is de GVG gekozen.



Figuur 2: Overeenkomst van GHG en GLG met de GVG. Alle grondwaterstanden in centimeters ten opzichte van het maaiveld (cm +mv). Gegevens: Runhaar (1989a&b).

De grondwaterstanden uit Schaffers & Sýkora (2000) zijn 17 keer gemeten (8 keer per jaar, van januari 1993 tot januari 1995). De gemiddelde voorjaarsgrondwaterstand (GVG) is berekend als het gemiddelde over de maanden maart en april. Een voor grondwaterstanden korte meetreeks van 2 jaar maakt een nadere beschouwing van de betrouwbaarheid van de gemiddelde grondwaterstanden nodig.

Volgens Van de Wouw (2000) is zelfs de GVG gemeten over 8 jaar is sterk beïnvloedbaar door het weer en dus niet heel erg betrouwbaar. Knotters & Bierkens (1999) berekenden dat de responstijden van de grondwaterstanden op neerslag sterk kunnen variëren (van enkele weken tot enkele jaren, hoewel de ondiepe grondwaterstanden vaak de kortste responstijden hebben). Er ontstaan systematische verschillen in de berekende GVG tussen plekken met verschillende responstijden:

dergelijke plekken zijn dan niet goed vergelijkbaar. Bij een GVG die gebaseerd is op een meetreeks van 2 jaar is dit effect nog sterker. Uit gegevens over het weer in en voorafgaand aan de meetperiode van Schaffers & Sýkora (2000) blijkt dat 1991 een jaar met weinig neerslag was, 1992 normaal, terwijl 1993 en 1994 (de jaren waarin gemeten is) respectievelijk een nat en een zeer nat jaar waren (KNMI 1990-1994). Daarnaast zijn de bemonsterde bermen verspreid over heel Nederland. De ruimtelijke variatie van de jaarlijkse neerslagsommen is erg groot (in 1994 bijvoorbeeld was de neerslagsom in De Bilt 220 mm hoger dan normaal, in Maastricht daarentegen 58 mm lager (KNMI 1994)). Dit maakt het erg onwaarschijnlijk dat een meetreeks van twee jaar een representatief beeld geeft van de grondwaterstanden. De grondwaterstandsgegevens van Schaffers & Sýkora zullen daarom niet gebruikt worden.

Ook bij de opnamen van Runhaar zijn gegevens over de grondwaterstand beschikbaar. De grondwaterstand is twee keer per maand gemeten in de periode 1980 – 1987. De GVG is weer het gemiddelde van de metingen in maart en april. De grondwaterstanden zijn over een periode van 8 jaar gemeten en de gemiddelden zullen vooral bij de ondiepe grondwaterstanden redelijk betrouwbaar zijn (zie echter hierboven). Na weglaten van de metingen die als 'onbetrouwbaar' aangegeven waren en opnamen waarvan de grondwaterstanden misten bleven 188 opnamen over. Om de beste weegmethode uit te zoeken zijn deze opnamegegevens en GVG's gebruikt.

2.2.3 Zuurgraad

Om de zuurgraad-indicatie te correleren met milieuvariabelen hebben Schaffers & Sýkora een aantal factoren gemeten. Omdat Schaffers & Sýkora vonden dat het *Reaktionszahl* van Ellenberg beter overeenkomt met de totale hoeveelheid calcium in de bodem dan met pH, zijn de gemiddelde indicatiewaarden met beide factoren vergeleken. De pH-CaCl₂ is per opname in januari en augustus gemeten. De totale hoeveelheid calcium is de som van de hoeveelheid uitwisselbaar Ca²⁺ en het Ca²⁺ dat aanwezig is in de vorm van calciumcarbonaat. Deze metingen zijn uitgevoerd in de bovenste 10 cm van de bodem.

Na de correlatieberekeningen in dit onderzoek bleek de correlatie van gemiddelde indicatiewaarde met het totale calciumgehalte juist lager te zijn dan de correlatie met pH. Volgens Schaffers & Sýkora (2000) was de correlatie tussen ongewogen gemiddelde indicatiewaarde en pH 0,79 en tussen de gemiddelde indicatiewaarde en het totale calciumgehalte (log-getransformeerd) 0,90. Waarschijnlijk is hiervoor de Pearson correlatiecoëfficiënt (r) gebruikt. Een eigen analyse van de correlatie met de Pearson correlatiecoëfficiënt met behulp van het statistisch pakket SPSS leverde dezelfde waarden op als de waarden die gepubliceerd zijn door Schaffers & Sýkora. In dit onderzoek is echter de Spearman rangcorrelatiecoëfficiënt (r_s) gebruikt. Voor de correlatie van de (volgens de kwalitatieve methode berekende) gemiddelde indicatiewaarde met het totale calciumgehalte werd een r_s van 0,815 berekend, de correlatie van de indicatiewaarden met de pH resulteerde in een r_s van 0,830. De rangcorrelatie met pH is dus groter dan de rangcorrelatie met het totale

calciumgehalte. Ook wanneer slechts de opnamen waarvan de pH hoger dan 5,25 is gebruikt worden, blijft de rangcorrelatie van gemiddelde indicatiewaarden met pH ($r_s = 0,401$) groter dan de rangcorrelatie met het totale calciumgehalte ($r_s = 0,360$).

Dit verschil in conclusies wanneer de Spearman rangcorrelatiecoëfficiënt dan wel de Pearson product-moment correlatiecoëfficiënt gebruikt wordt, kan waarschijnlijk teruggevoerd worden op de verdeling van de variabelen en de schalings- en transformatie-effecten. Ten eerste mag de Pearson correlatie slechts dan worden berekend, wanneer beide te correleren variabelen normaal verdeeld zijn en er een lineaire samenhang bestaat (zie o.a. Sachs 1984, maar in vele andere statistische handboeken zal deze informatie te vinden zijn). Wanneer met de toets van Kolmogorov-Smirnov nagegaan wordt of de variabelen wel normaal verdeeld zijn, blijkt dat alleen voor het log-getransformeerde calciumgehalte de aanname van normaliteit gerechtvaardigd is ($p = 0,151$). Bij de gemiddelde indicatiewaarden en de pH is er voldoende bewijs om deze aanname te verwerpen ($p = 0,022$ resp. $p = 0,000$). Ten tweede is de Pearson correlatiecoëfficiënt gevoelig voor de schaal waarop de waarnemingen liggen (vooral voor de aanwezigheid van 'gaten' en uitschieters in de gegevens) en de transformaties die uitgevoerd zijn, terwijl de rangcorrelatiecoëfficiënt daarentegen invariant is voor monotone transformaties (zie ook paragraaf 2.4.1). Het gevaar bestaat, dat bij de op normale verdelingen gebaseerde toetsen de conclusies veranderd worden door een transformatie uit te voeren. Wie er bijvoorbeeld voor zou kiezen om de gemiddelde indicatiewaarden te transformeren met een machtsfunctie, krijgt direct een ander beeld. Deze overwegingen geven genoeg grond om de Spearman rangcorrelatiecoëfficiënt te gebruiken in plaats van de Pearson correlatiecoëfficiënt. Omdat met de Spearman rangcorrelatiecoëfficiënt een hogere correlatie wordt gevonden tussen gemiddelde indicatiewaarde en pH dan tussen gemiddelde indicatiewaarde en het totale calciumgehalte, zal hier slechts de pH als milieufactor gebruikt worden.

De gegevens van Runhaar bevatten slechts metingen van de pH en het calciumgehalte van het grondwater. Omdat de zuurgraad en mineralenhuishouding van de wortelzone niet alleen door het grondwater worden bepaald, zijn deze metingen niet bruikbaar voor een correlatie van de indicatie van de vegetatie met het milieu.

2.3 Berekening van indicatiewaarden per opname

2.3.1 Bedekkingsweegwaarden

De bedekkingsweegwaarden die alle te onderzoeken transformaties opleveren kunnen ingevuld worden in formule (1), waarmee de gewogen gemiddelde indicatiewaarde van een opnamen berekend wordt. Voor de kwalitatieve methode zijn alle bedekkingsweegwaarden even groot. De weegwaarden van de kwantitatieve en decimale methoden moeten bepaald worden uit de schattingen van de bedekking van de soorten in de opnamen.

Omdat de bedekking, zoals die in het veld geschat is, meestal weergegeven is als klassen op een schaal, dient voor de kwantitatieve methode eerst het werkelijke bedekkingspercentage weer bepaald te worden uit de schattingscode. Wanneer de schattingscode naast bedekking in nauwe zin ook betrekking heeft op abundantie, wordt het toekennen van een bedekkingspercentage aan een schattingscode arbitrair. Hoe dit gedaan is voor de Braun-Blanquet schaal, waarmee alle opnamen die gebruikt zijn voor dit onderzoek gemaakt zijn, is te zien in Tabel 4.

Bij de decimale methode worden de cijfers 1 tot en met 9 aan de codes van de schaal van Braun-Blanquet toegekend (Tabel 5). De bedekkingsweegwaarde van de soort is dan het cijfer behorende bij de schattingswaarde. Het effect van deze transformatie is, dat de soorten met lage bedekking een relatief hoger gewicht krijgen. Deze berekeningswijze houdt ongeveer het midden tussen de kwantitatieve en de kwalitatieve methode.

Tabel 4: Bedekkingspercentages (naar Stevers et al. 1987) bepaald uit de Braun-Blanquet schaal

bedekking in het veld	abundantie in het veld	Braun-Blanquet code	bedekkingspercentage	opmerking
< 5%	zeer weinig individuen	r	0,1	bedekkingspercentage is arbitrair bepaald
< 5%	weinig individuen	+	0,3	
< 5%	veel individuen	1	1	
< 5%	zeer veel individuen	2m	3	
5 – 12,5%	n.v.t.	2a	8,75	bedekkingspercentage is het midden tussen boven- en ondergrens
12,5 – 25%	n.v.t.	2b	18,75	
25 – 50%	n.v.t.	3	37,5	
50 – 75%	n.v.t.	4	62,5	
75 – 100%	n.v.t.	5	87,5	

Tabel 5: Toekenning van de cijfers (gewichten) 1 – 9 aan de schaal van Braun-Blanquet

schattingswaarde	r	+	1	2m	2a	2b	3	4	5
cijfer	1	2	3	4	5	6	7	8	9

Ellenberg (1992) acht het gebruik van de kwantitatieve methode niet juist. Want, zo redeneert hij, de bedekking van een soort hangt vaak sterk af van de specifieke groeivorm van de soort. Grassen kunnen zich bijvoorbeeld gemakkelijk vegetatief uitbreiden, terwijl het zelfs onder optimale omstandigheden voor orchideeën moeilijk is om hoge bedekkingen te bereiken, omdat elk nieuw exemplaar zelf moet kiemen en groeien. De kwantitatieve methode zou alleen zin hebben bij soortenarme vegetaties, waar toevallige soorten anders een te groot gewicht zouden krijgen. Böcker et al. (1983) vonden kleine verschillen tussen een methode met bedekkingsafhankelijke gewichten en de kwalitatieve methode wanneer indicatiewaarden voor plantengemeenschappen uit de classificatie van Oberdorfer berekend werden. Als gewicht voor de indicatiewaarde van elke soort gebruikten Böcker et al. meestal het aantal keer dat de soort voorkwam in opnamen van een plantensociologische gemeenschap. Kowarik & Seidling (1989) berekenden de verschillen tussen de kwantitatieve en kwalitatieve methode voor vegetatieopnamen. De verschillen waren

voor enkele vegetatieopnamen soms vrij groot, vooral bij de N- en R-waarden volgens Ellenberg en bij anthropogene en extreem droge of natte vegetaties. Kowarik & Seidling bekeken echter niet welke methode de verschillen in het milieu het beste weergeeft.

De hiervoor genoemde weegwaarden zijn slechts enkele specifieke wijzen waarop de bedekking als gewicht in het berekenen van een gemiddelde indicatiewaarde kan worden meegenomen. Meer algemeen zou het bedekkingspercentage getransformeerd kunnen worden tot een bedekkingsweegwaarde (BWW) volgens een bepaalde functie:

$$BWW_i = f(B_i) \quad (2)$$

waarin B_i het bedekkingsaandeel of -percentage van soort i is.

Het is wenselijk dat de functie $f(B_i)$ zoveel mogelijk realistische transformaties uit kan voeren. Flip Witte (ongepubliceerd) heeft voor aanvang van dit onderzoek een formule toegepast die veel mogelijkheden om gewichten toe te kennen aan soorten, afhankelijk van hun bedekking, kan beschrijven:

$$BWW_i = a_0 + (1 - a_0) \left(\frac{B_i - B_{\min}}{B_{\max} - B_{\min}} \right)^{a_1} \quad (3)$$

met a_0 en a_1 als parameters die gevarieerd kunnen worden om de vorm van de grafiek van de functie te bepalen, en B_{\min} en B_{\max} als respectievelijk de minimale en maximale bedekking die mogelijk is volgens de Braun-Blanquet schaal (resp. 0,1% en 87,5%, volgens Tabel 4). De breuk $(B_i - B_{\min}) / (B_{\max} - B_{\min})$ schaalde de bedekking B_i naar een waarde tussen 0 en 1. Door deze breuk met $(1 - a_0)$ te vermenigvuldigen is de bedekkingsweegwaarde altijd gelijk aan 1 als $B_i = B_{\max}$. Op deze manier wordt voorkomen dat er functies ontstaan die de verschillende bedekkingen hetzelfde relatieve gewicht geven.

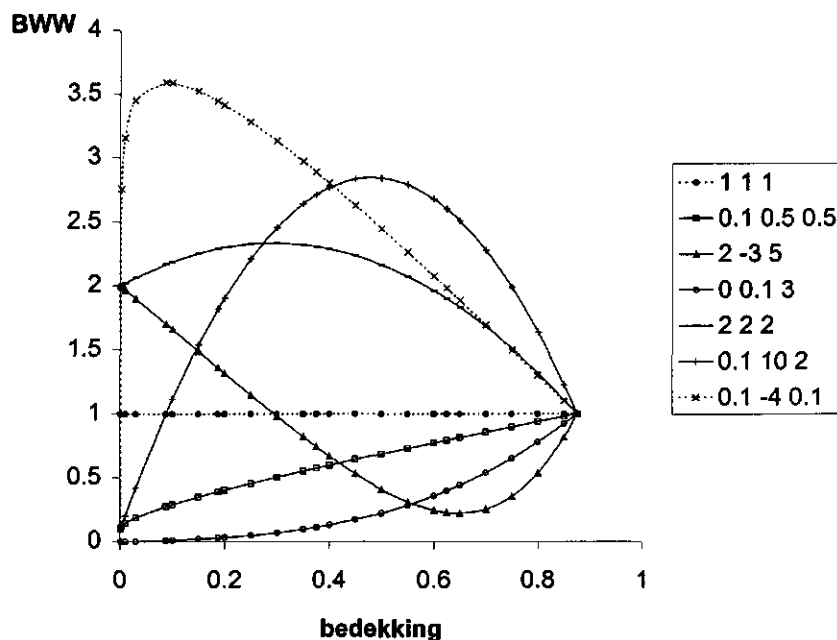
Functie (3) kan alleen een maximum of minimum bij $B_i = B_{\min}$ of $B_i = B_{\max}$ hebben, en kan dus slechts monotoon stijgende of dalende grafieken beschrijven. Een nadere beschouwing van het probleem maakt het wenselijk om meer verschillende curven te kunnen beschrijven. Soorten die vaak hoge bedekkingen bereiken kunnen soorten zijn met een hoge mate van ecologische tolerantie voor één of enkele milieufactoren, en dus weinig indicatief zijn. Soorten die erg weinig bedekken kunnen soorten zijn die toevallig in de opname aanwezig zijn, en 'verkeerde' informatie verschaffen over het milieu. Soorten met intermediaire bedekkingen zijn dan het meest indicatief. Die

soorten moeten dan, volgens deze redenering, ook het hoogste gewicht krijgen, en de soorten met heel hoge of lage bedekkingen een laag gewicht. Een bedekkingsweegwaardefunctie die deze vorm kan beschrijven wordt verkregen door een lineaire term aan functie (3) toe te voegen.

Het functievoorschrift verandert dan in:

$$BWW_i = a_0 + a_1 \left(\frac{B_i - B_{\min}}{B_{\max} - B_{\min}} \right) + (1 - a_0 - a_1) \left(\frac{B_i - B_{\min}}{B_{\max} - B_{\min}} \right)^{a_2} \quad (4)$$

Door de waarden van de parameters te variëren kan een groot aantal methoden die nooit eerder gebruikt zijn onderzocht worden. Enkele voorbeelden staan in Figuur 3. Ook de bestaande methoden kunnen met deze functie beschreven worden. Om de kwalitatieve methode te beschrijven moet $a_0 = 1$ en $a_1 = 0$, waarbij a_2 vrij is, of $a_0 = a_1 = a_2 = 1$. Voor de kwantitatieve methode dient $a_0 = 0$ en $a_1 = 1$ gekozen te worden, voor elke waarde van a_2 . De decimale methode is te benaderen met $a_0 = 0,11$, $a_1 = 0,03$ en $a_2 = 0,3$.



Figuur 3: Enkele voorbeelden van de omzetting van bedekkingen in gewichten zoals dat mogelijk is met formule (4). In de legenda rechts zijn de waarden van de parameters a_0 , a_1 en a_2 weergegeven.

De optimale bedekkingsweegwaardefunctie is te vinden door na te gaan welke combinatie van parameters de hoogste correlatie met een gemeten milieufactor oplevert. De waarden van de parameters variëren in discrete stappen. Alle mogelijke

combinaties van deze parameters zijn onderzocht, behalve als de combinatie van parameters negatieve weegwaarden mogelijk maakte. Om zoveel mogelijke relevante BWW-functies door te rekenen en tegelijk de rekentijd binnen de perken te houden, is gekozen om de stapgrootte afhankelijk te maken van de waarde van de parameters: in de buurt van de waarden 0 en 1 bevinden zich veel kleine stappen, hoe meer de waarden van 0 of 1 afwijken, des te groter de stapgrootte is. De waarden waarmee gerekend is staan in Tabel 6. Waar dit nodig bleek, zijn de intervallen verkleind om een nauwkeuriger schatting van de optimale waarden van de parameters te krijgen (tussen $a_0 = 0,0$ en $a_0 = 0,5$).

Tabel 6: De onderzochte waarden van de parameters in BWW-functie (4).

$a_0 =$	0.000	0.189	0.414	0.682	1.000	1.378	1.828	2.364	3.000	3.757	4.657	5.727	7.000
$a_1 =$	-5.038	-2.915	-1.493	-0.630	-0.187	-0.023	0.000	0.023	0.187	0.630	1.493	2.915	5.038
$a_2 =$	0.000	0.189	0.414	0.682	1.000	1.378	1.828	2.364	3.000				

Wanneer er behalve optimale parameters voor de BWW-functie ook optimale parameters voor andere weegmethoden (de hierna beschreven IWW- en SWW-functies) gevonden moesten worden, is voor elke combinatie van parameters van de andere weegmethode gezocht naar optimale parameters voor de BWW-functie.

2.3.2 Soortspecifieke bedekkingsweegwaarden

Om aan de kritiek tegemoet te komen dat sommige soorten makkelijker hoge bedekkingen bereiken dan andere, ook onder niet optimale omstandigheden, kan rekening gehouden worden met de soortspecifieke bedekking: de weegwaarde wordt berekend door de bedekking van elke soort in de opname te delen door een maat voor de soortspecifieke bedekking. Een dergelijke maat kan bijvoorbeeld de gemiddelde of maximale bedekking van een soort zijn. De bedekking van een soort wordt dan geschaald naar een voor die soort relevant gewicht. Een orchidee krijgt dan bij bijvoorbeeld 10% bedekking hetzelfde gewicht als sommige grassen bij 90%. Zo wordt voorkomen dat veeleisende soorten, die nooit hoge bedekkingen bereiken maar wel een smalle ecologische amplitude hebben, overschaduwd worden door snel groeiende soorten met een grote tolerantie voor milieuomstandigheden. Het potentiële voordeel van de kwantitatieve methode (wanneer er vanuit gegaan wordt dat hogere bedekkingen er op duiden dat de milieuomstandigheden sterker overeen komen met het optimum van de soort) blijft behouden.

Voor zover bekend is dit concept slechts één keer eerder toegepast, door Noest et al. (1989), voor clustering en ordinatie. Zij deelden de soorten in verschillende categorieën in, waarbij elke categorie een eigen bedekkings-drempelwaarde had. Kwam de bedekking in een opname boven de drempelwaarde, dan kreeg de soort het gewicht 9, als de bedekking onder de drempelwaarde bleef werd aan de soort het gewicht 1 toegekend. Dit systeem is weinig flexibel. Ook het gebruik van een drempelwaarde is niet erg elegant, omdat een bedekking die net iets hoger is dan de

drempelwaarde een disproportioneel groot gewicht krijgt ten opzichte van een bedekking net onder de drempelwaarde. Omdat bedekkingen slechts geschat zijn en afwijkingen van de werkelijkheid zeer waarschijnlijk zijn (zie bijvoorbeeld Lepš & Hadincová 1992) gaat de voorkeur uit naar een 'gladde' transformatiefunctie.

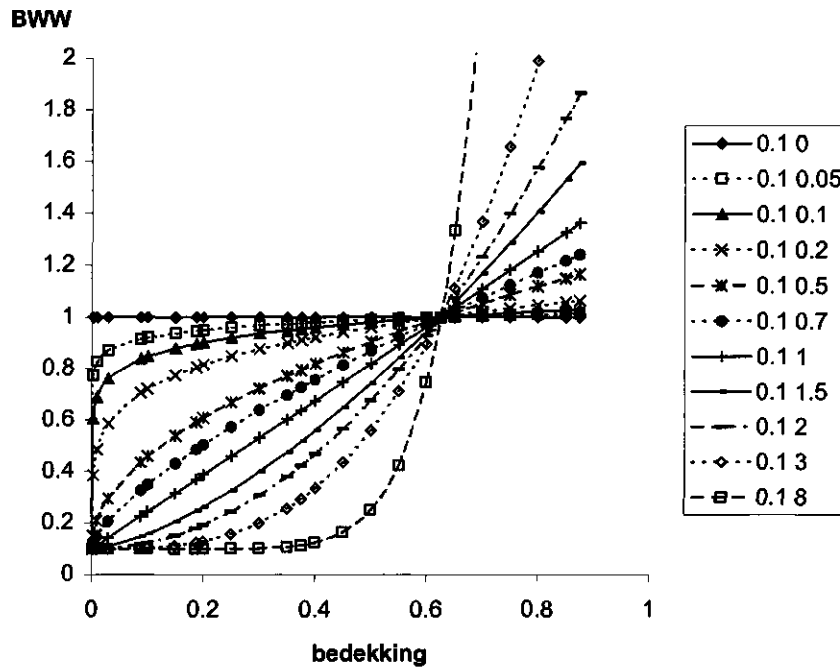
Omdat het voldoende is dat de bedekkingsweegwaardefunctie voor het rekenen met soortspecifieke bedekkingen monotone grafieken beschrijft, is er in het functievoorschrift geen lineaire term nodig. De functie die gebruikt is om de soortspecifieke bedekking mee te nemen in het berekenen van een bedekkingsweegwaarde is (analoog aan formule (3))

$$BWW_i = b_0 + (1 - b_0) \left(\frac{B_i - B_{\min}}{SSB_i - B_{\min}} \right)^{b_1} \quad (5)$$

waarin SSB_i de specifieke bedekking van soort i is. Hiervoor kan dus de gemiddelde bedekking of de gemiddelde maximale bedekking ingevuld worden.

Door b_0 (met stappen van 0,1) te variëren tussen 0 en 1 zijn alleen maar monotoon stijgende of horizontale (voor $b_0 = 1$) functies mogelijk. Dalende functies of niet-monotone functies zijn immers niet relevant. Enkele voorbeelden van transformaties van de bedekking in een bedekkingsweegwaarde met deze formule staan in Figuur 4.

Om de specifieke bedekking per soort te berekenen zijn de bedekkingsgegevens van ongeveer 35000 vegetatieopnamen die verzameld waren voor 'De Vegetatie van Nederland' (Schaminée et al. 1995) ter beschikking gesteld door Stefan Hennekens van Alterra. Er is gerekend met de gemiddelde bedekking per soort en de gemiddeld maximale bedekking (GMB). De gemiddelde maximale bedekking is bepaald door van elke soort de hoogste bedekkingen te nemen en deze te middelen. Hiervoor zijn de 25% (GMB 25%) of 10% (GMB 10%) hoogste bedekkingen genomen.



Figuur 4: Transformaties van de bedekking met formule (5). SSB_i is hier in alle gevallen 0,625. De waarden van b_0 en b_1 voor iedere grafiek staan vermeld in de legenda rechts naast de figuur ($b_0 = 0,1$ voor alle grafieken).

2.3.3 Ellenberg-indicatiewaarden en indicatieweegwaarden

Gebruikt zijn de Ellenberg-getallen (zoals de indicatiewaarden van Ellenberg vaak genoemd worden) voor vocht (F) en zuurgraad (R) zoals die vermeld zijn in de publicatie van Ellenberg et al. uit 1992. Per milieufactor kan een soort ingedeeld zijn in maximaal één klasse (indifferent of de waarden 1 tot en met 9 of 12), of de indicatiewaarde is onbekend. De soorten zonder indicatiewaarden zijn buiten beschouwing gelaten; de gemiddelde indicatiewaarde is berekend alsof deze soorten niet in de opname aanwezig waren. Indifferenten zijn óf op dezelfde manier als soorten zonder indicatiewaarden behandeld, óf volgens de methode die hieronder beschreven is.

Indifferente soorten zijn over het algemeen soorten met een te brede ecologische amplitude om ze in te kunnen delen bij een indicatiewaarde. Om in veel verschillende milieus te kunnen overleven, kunnen dergelijke soorten geen sterke aanpassingen aan extreme omstandigheden hebben, omdat aanpassingen aan extreme omstandigheden de concurrentiekracht van de soort in andere milieus aantasten. Een soort zonder aanpassingen aan extreme omstandigheden is dus gebonden aan niet-extreme milieus. Indifferenten komen dan meer in gemiddelde milieus voor dan in extreme. Met deze hypothese kunnen ook de indifferenten meegenomen worden bij het bepalen van de gemiddelde indicatiewaarde van de opname.

Het is mogelijk om indifferenten een gemiddelde indicatiewaarde (bij Ellenberg R: 5, en bij F: 6,5) en een laag gewicht te geven (bijvoorbeeld $BWW_i \times 0,1$). Kunzmann

(1989) stelt voor om indifferenten en de soorten met vochtgetal 5 samen te voegen. In een reactie hierop schrijft Ellenberg (1992, pag. 36) dat het verschil tussen indifferenten en soorten met vochtgetal 5 volgens hem wel gerechtvaardigd is. Hiernaast is het nadeel van een dergelijke behandeling van de indifferenten een verdere verkorting van de schaal waarop de berekende gemiddelde indicatiewaarden per opname liggen; het 'indicatiebereik' neemt nog verder af dan bij het berekenen van gemiddelden al het geval was.

Een andere mogelijkheid is het geven van meer gewicht aan de extreme indicatiewaarden naarmate er minder indifferenten zijn: indicatiewaarde-
weegwaarden, kortweg indicatieweegwaarden (*IWW*). Als er veel indifferenten in de opname aanwezig zijn krijgen de extreme indicatiewaarden een laag gewicht, als er weinig indifferenten zijn een groot gewicht. Extreme indicatiewaarden bij vocht zijn de hele 'natte' en 'droge' waarden, bij zuurgraad de erg 'zure' en 'basische' indicatiewaarden: de uiterste indicatiewaarden, die dus het verst van de middelste indicatiewaarden verwijderd zijn. De middelste indicatiewaarde is bij vocht 6,5 en bij zuurgraad 5. Het functievoorschrift moet er ongeveer als volgt uitzien:

$$IWW_I = 1 + \text{parameter} \times \left(\begin{array}{l} \text{term die groot is bij weinig} \\ \text{indifferenten en klein bij} \\ \text{veel indifferenten} \end{array} \right) \times \left(\begin{array}{l} \text{afstand van de} \\ \text{indicatiewaarde } I \\ \text{tot de middelste} \\ \text{indicatiewaarde} \end{array} \right) \quad (6)$$

Een mogelijkheid is:

$$IWW_I = 1 + c_1 \times \left(1 - \frac{c_2 \left(\sum BWW_{\text{indif}} \right)}{1 + c_2 \left(\sum BWW_{\text{indif}} \right)} \right) \times \left(\frac{I_c - I}{I_c - I_{\min}} \right)^2 \quad (7)$$

met:

c_1, c_2 = vrij te kiezen parameters;

I = indicatiewaarde, $1 \leq I \leq 9$ of $1 \leq I \leq 12$;

I_{\min} = kleinste mogelijke indicatiewaarde (= 1);

I_c = middelste indicatiewaarde, $I_c = 5$ of $I_c = 6,5$.

Deze indicatieweegwaardefunctie geeft de middelste indicatiewaarde het gewicht 1 en de meeste extreme indicatiewaarden een groter gewicht, afhankelijk van het aandeel indifferenten. Hoe meer indifferenten, des te lager het relatieve gewicht van de extreme indicatiewaarden. De parameter c_1 kan gevarieerd worden om het relatieve gewicht van de extreme indicatiewaarden ten opzichte van de middelste waarde te bepalen. De parameter c_2 kan gebruikt worden om het gewicht van de indifferenten te

bepalen: hoe groter c_2 , des te meer verschil is er tussen opnamen zonder en met indifferenten. Het effect van verschillende parameters in opnamen met een verschillend aantal indifferenten is weergegeven in Figuur 5.

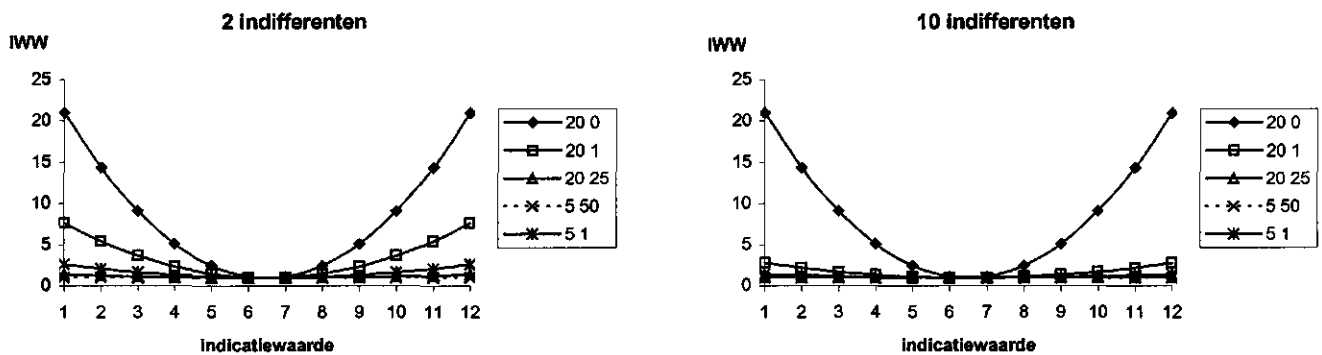
Voor het berekenen van de gemiddelde indicatiewaarde van een opname geldt nu:

$$I_{\text{gem}} = \frac{\sum_{i=1}^N I_i \times BWW_i \times IWW_i}{\sum_{i=1}^N BWW_i \times IWW_i} \quad (8)$$

waarin IWW_i de indicatieweegwaarde van de indicatiewaarde I_i van soort i is. Ook van formule (7) zijn de parameters geoptimaliseerd. De waarden van de parameters die onderzocht zijn staan in Tabel 7. Deze waarden zijn zo gekozen, dat de stapgrootte groter werd naarmate de waarde van de parameters meer afweek van 1.

Tabel 7: De onderzochte waarden van de parameters c_1 en c_2 in de IWW-formule (7).

$c_1 =$	0.000	0.704	0.963	1.000	1.037	1.296	2.000	3.370	5.630	9.000	13.704	19.963
	28.000	38.037	50.296	65.000								
$c_2 =$	0.000	0.704	0.963	1.000	1.037	1.296	2.000	3.370	5.630	9.000	13.704	19.963
	28.000	38.037	50.296	65.000								



Figuur 5: Effect van verschillende parameters c_1 en c_2 (waarden weergegeven in de legenda's rechts van de grafieken) en aantal indifferenten op de weegwaarden van de indicatiewaarden voor vocht volgens de IWW-formule (7). Er is voor kwalitatieve weging van de bedekking gekozen.

2.3.4 Ecotopensysteem en soortweegwaarden

2.3.4.1 Van ecologische groepen naar soorten

De indicatiewaarde van een soort voor een bepaalde milieufactor is af te leiden uit de soortengroep(-en) waarin de soort ingedeeld is. Voor vocht, voedselrijkdom en pH is de indicatie eenvoudig af te leiden uit de code van de soortengroep. Het eerste cijfer heeft betrekking op vocht, het tweede op zuurgraad en voedselrijkdom. Het toekennen van een indicatiewaarde aan elke klasse is weergegeven in Tabel 8 en Tabel 9.

Tabel 8: Toekenning van indicatiewaarden aan ecologische groepen: vocht.

1e cijfer	omschrijving	vocht-indicatiewaarde
1	aquatisch	1
2	nat	2
4	vochtig	3
6	droog	4

Tabel 9: Toekenning van indicatiewaarden aan ecologische soortengroepen: voedsel en zuurgraad.
X = indifferent, - = geen indicatie.

2e cijfer	omschrijving	zuurgraad-indicatiewaarde	voedsel-indicatiewaarde (niet gebruikt)
0		-	-
1	voedselarm zuur	1	1
2	voedselarm zwak zuur	2	1
3	voedselarm basisch	3	1
4	voedselarm	X	1
7	matig voedselrijk	-	2
8	zeer voedselrijk	-	3
9	matig tot zeer voedselrijk	-	2 en 3

Omdat één soort in meer soortengroepen ingedeeld kan zijn, moet hiervoor bij het berekenen van een gemiddelde indicatiewaarde gecorrigeerd worden. Vanwege het hierboven genoemde criterium wordt aangenomen dat elke soortengroep waar de soort in voorkomt een evenredig deel van het voorkomen van die soort verklaart (Witte 1998). De indicatiewaarde wordt dan gewogen met de fractie van de groepen die horen bij de indicatiewaarde van het totaal: elke indicatiewaarde heeft een eigen weegfactor (let op: dit is geen weegwaarde zoals gebruikt bij de indicatieweegwaarden).

De indicatiewaarde van soort i is dan:

$$I_i = \sum_{j=1}^M F_{i,j} \times j. \quad (9)$$

waarin $F_{i,j}$ de weegfactor van indicatiewaarde j van soort i is en M het aantal indicatiewaarden voor de betreffende milieufactor. Per soort geldt dat

$$\sum_{j=1}^M F_{i,j} = 1. \quad (10)$$

De gemiddelde indicatiewaarde van de opname wordt dan, door substitutie van vergelijking (9) in (1), berekend volgens:

$$I_{\text{gem}} = \frac{\sum_{i=1}^N \left(\sum_{j=1}^M F_{i,j} \times j \right) \times BWW_i}{\sum_{i=1}^N BWW_i} \quad (11)$$

Zoals te zien is in Tabel 8 en Tabel 9 komen indifferenten alleen voor bij de indicatiewaarden voor zuurgraad. Ook komen er soorten voor die wel voor de andere factoren een indicatiewaarden hebben, maar niet voor zuurgraad. Dit hangt samen met de keuze die er in het ecotopensysteem gemaakt is, namelijk om bij matig tot zeer voedselrijke soortengroepen geen onderscheid te maken naar zuurgraad. Deze soorten hebben dus geen indicatie voor zuurgraad. De groepen van brakke en zoute bodems hebben voor zowel voedselrijkdom als zuurgraad geen indicatie (bijvoorbeeld bP20). Er zijn drie ecotopen (R24, R44 en R64) van voedselarme bodem, waarbij geen onderscheid gemaakt is naar zuurgraad. De soorten uit deze ecotopen zijn als indifferent beschouwd.

Het kan voorkomen dat een soort naast de indeling in één of meerdere zuurindicatieklassen ook in indifferente soortengroepen of in soortengroepen zonder zuurindicatie is ingedeeld (zie bijvoorbeeld *Molinia caerulea*, Tabel 10). Ook zijn er enkele soorten die in zowel zure, zwak zure als basische milieus voorkomen, maar niet als indifferent geïnclassificeerd zijn omdat ze niet in ruigten voorkomen (en voedselarme, indifferente soortengroepen niet bestaan buiten ruigten). Omdat het als indifferent geïnclassificeerd zijn van soorten samen hangt met de indicatie voor structuur, is een indicatieweegwaarde gebaseerd op het aandeel indifferenten van de

opnamen hier niet zinnig. De soorten die voor meer dan de helft in matig tot zeer voedselrijke ecotopen en ruigten voorkomen zijn bij het berekenen van de gemiddelde indicatiewaarde voor zuurgraad geheel buiten beschouwing gelaten.

Van de soorten die overbleven is bij het berekenen van de gemiddelde indicatiewaarde voor zuurgraad volgens formule (9) alleen gebruik gemaakt van de indicatieklassen 1 t/m 3 en de bijbehorende weegfactoren, waarbij de weegfactoren gecorrigeerd zijn voor een eventuele toedeling van soorten aan klassen 'indifferent' en 'geen indicatie'. Als bijvoorbeeld een soort voor 2/3 in basische ecotopen voorkomt en voor 1/3 in matig voedselrijke, dan krijgt deze soort voor formule (9) de indicatiewaarde 3 met weegfactor 1.

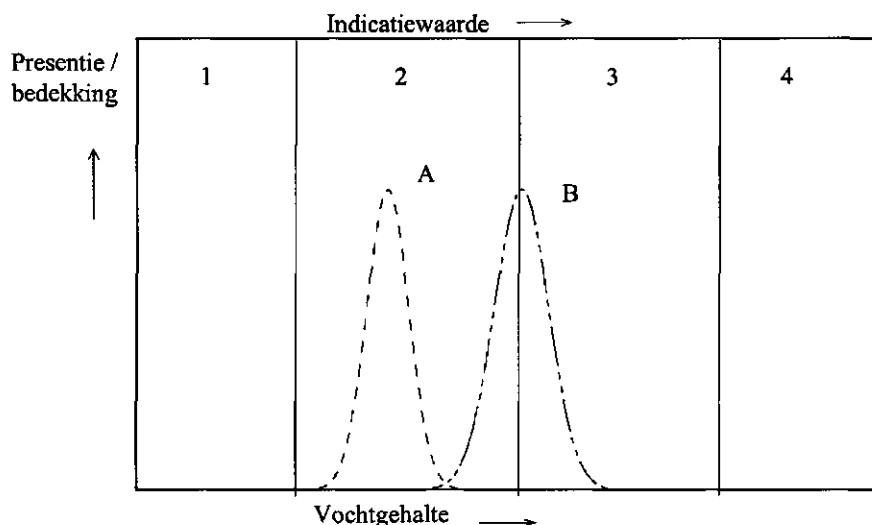
Een voorbeeld is te vinden in Tabel 10. De vocht-indicatiewaarde voor *Molinia caerulea* is dan 2,5; de zuurgraad-indicatiewaarde is 1,5 (de weegfactoren voor $j = 1, j = 2$ en $j = 3$ zijn herschaald tot $\sum_{j=1}^3 F_j = 1$) en de voedsel-indicatiewaarde is 1.

Tabel 10: Toekenning van weegfactoren aan *Molinia caerulea*, die voorkomt in soortengroepen G21, G22, G41, G42, R24, R44

vocht		zuurgraad		voedselrijkdom	
indicatie	weegfactor	indicatie	weegfactor	indicatie	weegfactor
1	0/6 = 0	1	2/6 = 0,33	1	6/6 = 1
2	3/6 = 0,5	2	2/6 = 0,33	2	0/6 = 0
3	3/6 = 0,5	3	0/6 = 0	3	0/6 = 0
4	0/6 = 0	X	2/6 = 0,33		

2.3.4.2 Soortweegwaarden

Er zijn twee gevallen waarin een soort in meerdere indicatiewaardeklassen ingedeeld wordt. Het kan zijn dat de soort een brede amplitude heeft, maar het is ook mogelijk dat een soort met smalle amplitude het optimum dicht bij een klassengrens heeft. Dit laatste wordt geïllustreerd in Figuur 6.



Figuur 6: De verspreiding van twee soorten (A en B) met gelijke amplitude langs een vochtgradiënt en de indeling in klassen van het ecotopensysteem. Soort A wordt volledig ingedeeld bij indicatiewaarde 2, terwijl soort B bij indicatiewaarde 2 en 3 met weefactor 0,5 wordt ingedeeld.

Bij formule (9) wordt uitgegaan van gelijke amplituden van elke soort: de weefactoren worden slechts gebruikt om een gemiddelde indicatiewaarde per soort te berekenen. In bijvoorbeeld het geval van de vochtindicatie van *Molinia caerulea* maakt het ook niets uit of er met indicatiewaarde 2,5 en weefactor 1 gerekend wordt, of met de indicatiewaarden 2 en 3 waarbij beide indicatiewaarden weefactor 0,5 hebben.

Het is echter waarschijnlijk dat soorten die in meer dan één indicatieklasse voorkomen een bredere amplitude hebben dan soorten die slechts in één klasse ingedeeld zijn. Het voorbeeld in Figuur 6 (twee soorten met even brede amplitude en een verschillend breed indicatiespectrum) zal eerder uitzondering dan regel zijn, aangezien de klassen in het ecotopensysteem breed gedefinieerd zijn. Omdat door de indeling aangegeven wordt hoe breed de amplitude van de soort ongeveer is, biedt dit een uitstekende mogelijkheid om een amplitude-afhankelijk gewicht te geven aan een indicatiewaarde van de soort. Schaffers & Sykora (2000) gebruiken de standaardfout, die berekend is uit de tolerantie (afgeleid uit meetgegevens), als gewicht bij het berekenen van de correlatie tussen soortoptima en Ellenberg-indicatiewaarde. Runhaar et al. (1997) berekenden een hogere verklaarde variatie bij de correlatie van het aandeel natte of droge soorten met de vochttoestand van de bodem als ze de soorten die in meerdere vochtgroepen voorkwamen weglieten.

Een maat voor de breedte van de amplitude is het aantal indicatieklassen waarin een soort ingedeeld is. De indicatiewaarde kan worden gewogen met een maat voor de spreiding van de soort over de verschillende indicatieklassen: een soortweefwaarde (*SWW*). Een eenvoudige manier om dit te doen is

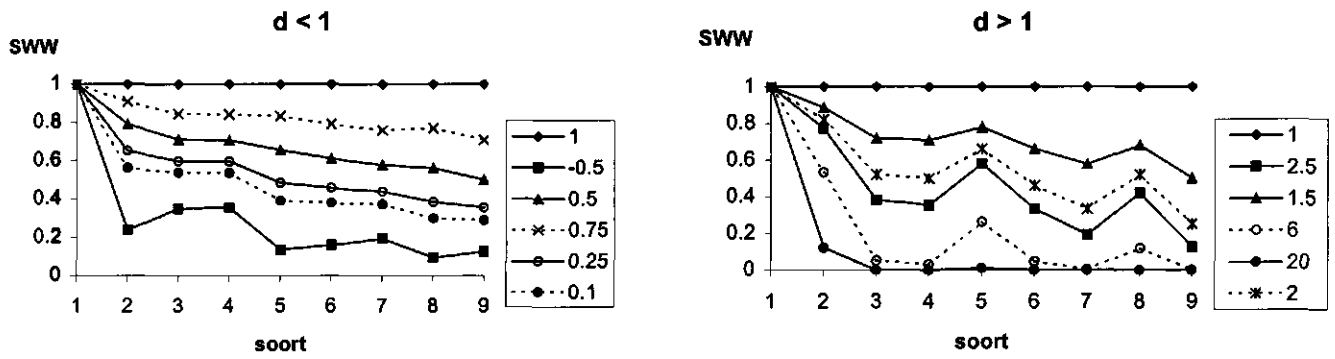
$$SWW_i = \sum_{j=1}^M F_{i,j}^d \quad \text{voor } d \geq 1. \quad (12)$$

Een soort die in drie indicatieklassen voorkomt en waarvan elke weegfactor 0,33 is, krijgt hierbij een lager gewicht dan een soort die in drie klassen voorkomt met de weegfactoren 0,6, 0,3 en 0,1, en deze soort krijgt weer een lager gewicht dan een soort die in twee klassen voorkomt met de weegfactoren 0,6 en 0,4. Soorten die een duidelijk optimum hebben, dat af te leiden is uit de verdeling in indicatieklassen, krijgen dus een hoog gewicht.

Iets andere eigenschappen heeft de soortweegwaardefunctie met d kleiner dan 1. Het geheel moet dan wel omgekeerd worden om ervoor te zorgen dat soorten met een smalle amplitude de hoogste weegwaarde krijgen. De SWW wordt op deze wijze veel sterker afhankelijk van het aantal weegfactoren dan de grootte van die factoren:

$$SWW_i = \frac{1}{\sum_{j=1}^M F_{i,j}^d} \quad \text{voor } 0 \leq d \leq 1. \quad (13)$$

Deze formule is veel gevoeliger voor het aantal weegwaarden (en dus het aantal klassen waarin een soort ingedeeld is) dan de functie (12). De functie (12) met $d = 2$ is eigenlijk het omgekeerde van Simpson's diversiteitsindex. Een voorbeeld: voor $d = 2$ en functie (12) geeft de set weegfactoren (0,7;0,1;0,1;0,1) als soortweegwaarde 0,56 en de set (0,33; 0,33; 0,33) geeft 0,58. Voor $d = 0,5$ en functie (13) levert dit respectievelijk 0,52 en 0,33 als soortweegwaarden op. Een illustratie van het effect van de verschillende waarden van de parameter d is te zien in Figuur 7. Soort 1 (met slechts 1 weegfactor) heeft de smalste amplitude, en dus de hoogste soortweegwaarde. De soortweegwaarden van de andere soorten verschillen, afhankelijk van de waarde van de parameter d . Door veel verschillende parameters te gebruiken kan op een objectieve manier bepaald worden welke waarde van d het beste resultaat geeft (zie Tabel 12 voor de gebruikte waarden van d). Of een dergelijke optimalisering van diversiteits-indices, waar deze weegwaarde veel van weg heeft, wel eens plaatsgevonden heeft is niet bekend.



Figuur 7: Het effect van verschillende waarden van de parameter d (weergegeven in de legenda's rechts naast de grafieken) in de SWW-formules (12) en (13) op de weegwaarden van soorten met verschillende amplituden en weegfactoren. Tabel 11 geeft de weegfactoren voor de soorten.

Tabel 11: De weegfactoren van de soorten die gebruikt zijn voor Figuur 7.

soort	weegfactoren ($F_{i,j}$)			
1	1			
2	0.9	0.1		
3	0.6	0.4		
4	0.5	0.5		
5	0.8	0.1	0.1	
6	0.6	0.3	0.1	
7	0.33	0.33	0.33	
8	0.7	0.1	0.1	0.1
9	0.25	0.25	0.25	0.25

Tabel 12: De onderzochte waarden van de parameter d in de SWW-formules (12) en (13).

$d =$	-1.370	-0.372	0.298	0.704	0.912	0.989	1.000	1.011	1.088	1.296	1.702	2.372
	3.370	4.764	6.619	9.000								

De uiteindelijke formule voor het berekenen van een gemiddelde indicatiewaarde van een opname ziet er als volgt uit (aanpassing van formule (11)):

$$I_{\text{gem}} = \frac{\sum_{i=1}^N \left(\sum_{j=1}^M F_{i,j} \times j \right) \times BWW_i \times SWW_i}{\sum_{i=1}^N BWW_i \times SWW_i} \quad (14)$$

2.3.5 Mediaan

Omdat de mediaan, berekend op gangbare wijze als de middelste waarneming, alleen bepaalde waarden aan kan nemen is deze niet geschikt voor een fijne afstemming zoals die hier gewenst is (Kowarik & Seidling 1989). O.a. Sachs (1984, pag. 75) en Lienert (1986, pag. 45) beschrijven een methode voor het berekenen van de mediaan als de waarnemingen in klassen ingedeeld zijn. Voor de berekening van de mediaan op deze wijze worden alle indicatiewaarden opgevat als klassen die lopen van $I - \frac{1}{2}$ tot $I + \frac{1}{2}$, voor elke indicatiewaarde I . De klasse waarin alle soorten met indicatiewaarde 3 zitten loopt bijvoorbeeld van 2,5 tot 3,5. De formule voor berekening van de mediaan luidt

$$I_m = O + b \times \left(\frac{(N/2) - N_o}{N_i} \right) \quad (15)$$

met

- I_m = mediane indicatiewaarde;
- O = de ondergrens van de klasse waarin de middelste waarneming valt;
- b = de breedte van de klasse waarin de middelste waarneming valt;
- N = het aantal soorten per opname;
- N_o = het aantal soorten onder de klasse waarin de middelste waarneming valt;
- N_i = het aantal soorten in dezelfde klasse als de middelste waarneming.

Bij berekening van de mediaan van een opname waarin de volgende Ellenberg indicatiewaarden voorkomen (1, 2, 3, 3, 4, 4, 4, 4, 4, 5, 5) geldt: $O = 3,5$; $b = 1$; $N = 11$; $N_o = 4$; $N_i = 5$ en $I_m = 3,8$ (volgens de gangbare berekening zou dit 4 geweest zijn). Met de indicatiewaarden (1, 3, 3, 3, 4, 4, 5, 6) is I_m precies gelijk aan 3,5 (net als met de gangbare methode).

Bij het ecotopensysteem wordt de indicatiewaarde van een soort bepaald als gewogen gemiddelde over de ecologische groepen. Zo zijn ook niet-gehele getallen mogelijk als indicatiewaarde, zodat de indicatiewaarden van de afzonderlijke soorten niet meer zonder meer in klassen ingedeeld zijn. De mediane indicatiewaarden zijn alleen bij het systeem van Ellenberg berekend.

2.4 Statistische verwerking

2.4.1 Correlatie

Voor elke serie gemiddelden of medianen wordt de Spearman rangcorrelatiecoëfficiënt (r_s) met correctie voor knopen (een knoop is een aantal waarnemingen met dezelfde waarde) (Lienert 1986, pag. 598) berekend als maat voor de correlatie tussen de gemiddelden per opname en de gemeten milieufactor. De formule voor het berekenen van r_s met knopencorrectie is

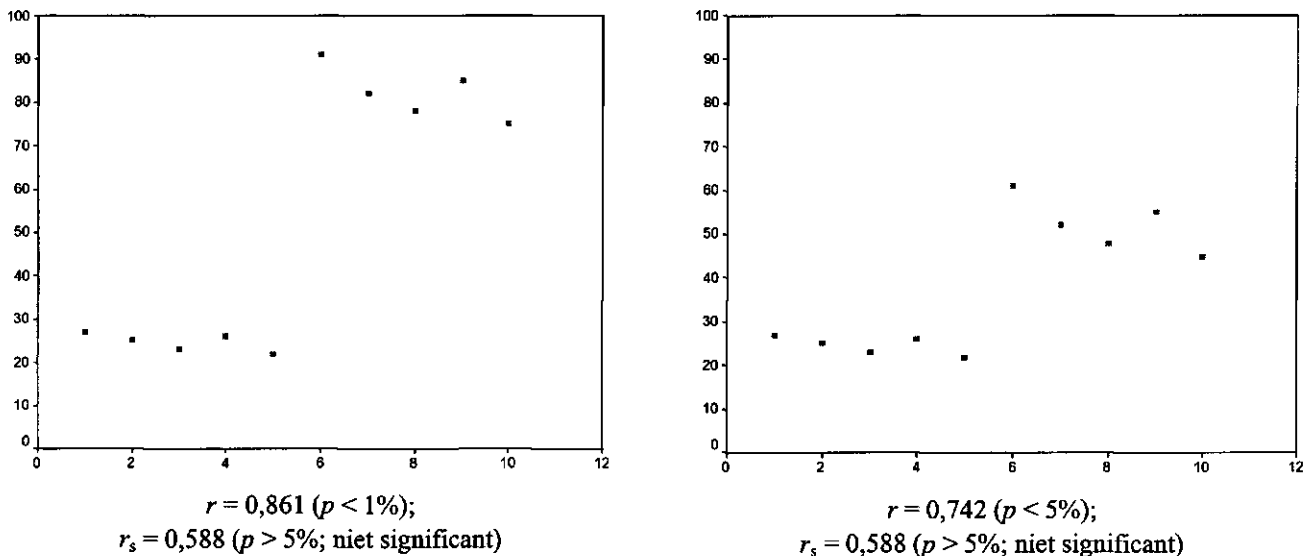
$$r_s = \frac{n^3 - n - \frac{1}{2} \left(\sum_{i=1}^K t_i^3 - t_i \right) - \frac{1}{2} \left(\sum_{j=1}^L u_j^3 - u_j \right) - 6 \sum_{m=1}^n d_m^2}{\sqrt{\left(n^3 - n - \left(\sum_{i=1}^K t_i^3 - t_i \right) \right) \left(n^3 - n - \left(\sum_{j=1}^L u_j^3 - u_j \right) \right)}} \quad (16)$$

met:

- n = aantal paren waarnemingen;
- K = aantal knopen bij de gemiddelde indicatiewaarden;
- t_i = grootte van knoop i ;
- L = aantal knopen bij de gemeten milieufactor;
- u_j = grootte van knoop j ;
- d_m = verschil tussen rangnummer van de m^{de} gemiddelde indicatiewaarde en rangnummer van de bijbehorende gemeten milieufactor.

Het voordeel van r_s boven de Pearson product-moment correlatie (r) is dat r_s onafhankelijk is van een (bijvoorbeeld log-) transformatie van de gegevens. Ook is r_s ongevoelig voor de lengte van de schaal waarop de waarnemingen liggen. Statistisch gezien is het gebruik van r alleen juist als de te correleren variabelen beide normaal verdeeld zijn en er een lineaire samenhang bestaat. r_s is daarentegen verdelingsvrij en mag bij elke monotoon verband toegepast worden.

De punten in Figuur 8 (beide grafieken) zijn alleen positief met elkaar gecorreleerd als geheel, binnen de groepen is er eerder een negatieve correlatie. Sachs (1984) noemt dit een "*Inhomogenitätskorrelation*". Zoals te zien is, beschrijft de Spearman rangcorrelatiecoëfficiënt deze situatie het beste.



Figuur 8: De Pearson correlatiecoëfficiënt (r) is gevoelig voor inhomogeniteit van de gegevens, de Spearman rangcorrelatiecoëfficiënt (r_s) veel minder. De conclusie gebaseerd op r_s (geen significante correlatie) is in beide gevallen de juiste.

2.4.2 Schatting van de correlatie met behulp van de bootstrap

De samenstelling van de steekproef die wordt gebruikt om de correlatiecoëfficiënt te berekenen is van invloed op de waarde van de correlatiecoëfficiënt (net als bij andere statistische schatters). Bij een andere steekproef was de waarde van de correlatiecoëfficiënt waarschijnlijk ook anders geweest. Het behoeft geen uitleg dat al te veel stochastische variatie in de correlatiecoëfficiënt ongewenst is, aangezien we er de werkelijke correlatie mee willen schatten. We zouden graag willen, dat de resultaten van elk onderzoek een algemene geldigheid hebben, en niet slechts gelden voor de steekproeven die toevallig in het onderzoek gebruikt zijn. Nu is het bij onderzoek aan de vegetatie niet mogelijk om een statistisch verantwoorde, aselechte steekproef te nemen, omdat op die manier slechts een klein deel van de soorten bemonsterd wordt (Witte & Runhaar 2000). De gegevens die gebruikt zijn in dit onderzoek zijn dus in zekere mate bepaald door de selectiviteit van steekproefname, en gedeeltelijk door toeval. Om een schatter van de correlatie te krijgen waarin het toeval nauwelijks een rol speelt, is de bootstrap-methode toegepast. Bootstrapping is helaas geen wondermiddel: de selectiviteit van steekproefname is ermee niet uit te vlakken. Toch is een voor toevalligheden gecorrigeerde schatter van de correlatie al een grote vooruitgang. Hiermee wordt het selecteren van de beste methode een stuk overzichtelijker.

Bootstrapping is gebaseerd op een op zich eenvoudig idee, en kan op zeer veel problemen toegepast worden (Efron & Tibshirani 1993). De bootstrap gaat uit van de gegevens, zonder aannames te maken over de verdeling, en is dus verdelingsvrij (Hjorth 1994). De algemene vorm van de bootstrap-methode is beschreven in Efron & Tibshirani (1993). Diaconis & Efron (1983) beschrijven de bootstrap op een heldere,

illustratieve wijze. Hier volstaat alleen de specifieke werkwijze voor het bepalen van een bootstrap-schatter van de rangcorrelatiecoëfficiënt.

Correlatie gaat uit van n in paren gegroepeerde waarnemingen, x_i en y_i ($i = 1, 2, 3, \dots, n$) vormen die paren. Alle x -waarnemingen vormen de vector \mathbf{x} , idem voor de vector \mathbf{y} . De rangcorrelatiecoëfficiënt tussen deze twee vectoren wordt dan berekend na rangnummers van de waarnemingen in elke vector. Voor de bootstrap-methode worden uit deze waarnemingen aselekt, met teruglegging, n nieuwe paren getrokken, waarbij de oorspronkelijke paren intact blijven. De nieuwe steekproef bestaat uit de vectoren \mathbf{x}^* en \mathbf{y}^* . Het is waarschijnlijk dat een aantal van de oorspronkelijke waarnemingen meerdere keren in deze nieuwe vectoren voorkomen, terwijl andere waarnemingen ontbreken. De correlatie tussen deze vectoren (r^*) wordt, na rangnummers van de waarnemingen, berekend met formule (16). Dit trekken van nieuwe vectoren en berekenen van r^* gebeurt een groot aantal keer. In dit onderzoek bleek 1000 herhalingen voldoende. De bootstrap-schatter van de correlatie is de mediaan van alle r^* . Alle hierna genoemde correlatiecoëfficiënten zijn geschat door middel van de bootstrap-methode.

2.4.3 Het vergelijken van twee correlatiecoëfficiënten met behulp van een permutatietest

Het is niet mogelijk om via een klassieke formule vast te stellen of het verschil tussen twee rangcorrelatiecoëfficiënten statistisch significant is (Lienert 1986 pag. 601). In deze studie is het zeer zeker niet mogelijk om de correlatiecoëfficiënten met een standaardformule te vergelijken (zelfs als die zou bestaan), omdat de coëfficiënten die vergeleken moeten worden gebaseerd zijn op dezelfde gegevens en dus niet onafhankelijk van elkaar zijn. Met een methode die ontwikkeld is om de correlatie van twee aselechte steekproeven met elkaar te vergelijken zal immers bijna nooit een significant verschil gevonden worden wanneer die steekproeven in sterke mate met elkaar samen hangen (zie het voorbeeld in Bijlage 5 voor het verschil tussen de hierna besproken permutatietest en de z -transformatie van de correlatiecoëfficiënt, een klassieke methode).

Ook hier echter biedt de computer een oplossing. Wanneer vergeleken moet worden welke van de methoden de hoogste correlaties van gemiddelde indicatiewaarden met het milieu geeft, zijn drie vectoren nodig: een vector met indicatiewaarden volgens methode 1 (\mathbf{w}), een vector met indicatiewaarden volgens methode 2 (\mathbf{x}) en een vector met de gemeten waarden van de milieufactor waarmee gecorreleerd wordt (\mathbf{y}). In de nu te gebruiken permutatietest (zie onder andere Efron & Tibshirani 1993, hoofdstuk 15) worden de paren waarnemingen uit de vectoren \mathbf{w} en \mathbf{x} willekeurig verwisseld, zodanig dat volledig gemengde nieuwe vectoren (\mathbf{w}^* en \mathbf{x}^*) ontstaan. Hiertoe dient de vector \mathbf{q} . De waarde van q_i bepaalt of een paar waarnemingen (w_i, x_i) omgewisseld wordt of niet. Een randomfunctie kent aan elke q_i ($i = 1, 2, 3, \dots, n$) de waarden 0 of 1 toe, beide met kans $\frac{1}{2}$. Bij de waarde 0 wordt het paar (w_i, x_i) niet omgewisseld, en bij 1 wel. Zo ontstaan twee nieuwe, gepermuteerde vectoren \mathbf{w}^* en \mathbf{x}^* . De

correlatiecoëfficiënten van beide vectoren met y (die onveranderd is) zijn respectievelijk r^*_w en r^*_x . Het verschil tussen r^*_w en r^*_x (hier s^* genoemd) wordt een groot aantal (K) keer berekend, waarbij q steeds opnieuw *random* gevuld wordt. Door de permutatie van w en x ontstaat er een verdeling van s^* op basis van de nulhypothese (er is geen verschil tussen beide methoden). De kans dat een s^* , als de nulhypothese waar is, groter is dan het verschil tussen de correlatiecoëfficiënten berekend uit de originele vectoren (s), is het aantal keer dat s^* daadwerkelijk groter is dan s , gedeeld door het aantal permutaties K . Nogmaals, in formulevorm:

$$p = \frac{\#(s^* \geq s)}{K} \quad (17)$$

waarin p de eenzijdige overschrijdingskans is. Omdat de verdeling van s^* nagenoeg symmetrisch is, is de tweezijdige overschrijdingskans eenvoudigweg twee maal de eenzijdige.

Voor het vergelijken van twee correlatiecoëfficiënten zijn in dit onderzoek steeds 1000 herhalingen gebruikt. Meer herhalingen leverde geen wezenlijk andere overschrijdingskansen op, hoewel de 1000 verschillende vectoren q slechts een fractie vormen van het totaal aantal mogelijkheden (bij 188 opnamen zijn dit 2^{188} mogelijkheden). Meer herhalingen kunnen een meer accurate schatting geven van zeer kleine overschrijdingskansen, maar voor dit onderzoek is het verschil tussen een overschrijdingskans van bijvoorbeeld 0,1 % en 0,3 % niet van belang. Ook deze permutatietest gebruikt alleen de waarnemingen, en is dus verdelingsvrij, en is te gebruiken bij niet aselekt genomen steekproeven (Edgington 1995).

Voor de daadwerkelijke permutatie (het al of niet verwisselen van de paren) van de vectoren w en x moeten de waarnemingen in deze vectoren ongeveer hetzelfde verdeeld zijn. Dit wordt bereikt door alle waarnemingen in de vectoren als volgt te schalen: van elke waarneming wordt het gemiddelde van alle waarnemingen in de vector afgetrokken, waarna de waarde wordt gedeeld door de standaardafwijking van dit gemiddelde. De waarnemingen in beide vectoren krijgen dan een verdeling met standaardafwijking 1 rond het gemiddelde 0. Wanneer dit achterwege blijft, betekent bijvoorbeeld een overschrijdingskans van 1,5% dat de verdelingen van beide vectoren verschillend zijn, en niet specifiek de correlatiecoëfficiënten.

Er zijn voor deze test twee soorten nul- en alternatieve hypothesen gebruikt. Wanneer er geen reden was om aan te nemen dat één van beide methoden beter is, is de nulhypothese 'geen verschil' tegen de alternatieve hypothese 'wel verschil' getest. Hiervoor dienen de tweezijdige overschrijdingskansen gebruikt te worden. Wanneer van tevoren vast staat welke methode de hoogste correlatiecoëfficiënt moet hebben (laten we deze methode 'A' noemen), is het triviaal om te testen of deze methode slechter is dan de andere. In dit geval is de nulhypothese 'geen verschil' tegen de

alternatieve hypothese 'methode A is beter' getest, door gebruik te maken van eenzijdige overschrijdingskansen. Dit laatste doet zich altijd voor bij het vergelijken van de optimale methode met bestaande methoden, of bij het vergelijken van een berekening mét een extra methode (bijvoorbeeld *IWW*) met een berekening zónder die extra methode.

Een klassieke test gaat er vanuit dat er een aselechte steekproef gebruikt wordt, en de uitspraak van zo'n test is een uitspraak over de populatie. Wanneer de steekproef niet aselekt is, is een uitspraak over de hele populatie op basis van de steekproef niet mogelijk en het gebruik van klassieke methoden dus verkeerd (Edgington 1995, meer hierover in paragraaf 4.1). De permutatietest laat het aan de gebruiker over om te bepalen of de steekproef representatief is voor de hele populatie. In het geval van een aselechte steekproef zal de uitkomst van een permutatietest zonder meer geldig zijn voor de hele populatie (net als een klassieke test). Aangezien de vegetatieopnamen in deze studie geen aselechte steekproef vormen, zal een significant resultaat in beginsel alleen significant zijn voor de specifieke dataset die gebruikt is. Als er redenen zijn om aan te nemen dat de steekproef niet representatief is, geldt het resultaat niet voor vegetatieopnamen in het algemeen.

3 Resultaten

3.1 Bedekkingsweegwaarden

3.1.1 Algemeen

De 188 opnamen van Runhaar die gebruikt zijn voor het vinden van de beste *BWW*-functie voor vochtindicatie bevatten gemiddeld 21,4 soorten per opname. Voor het gebruik van Ellenberg F-getallen blijven na verwijdering van de indifferenten en soorten zonder indicatiewaarde gemiddeld 17,8 soorten over voor de berekening, met een minimum van 3 soorten en een maximum van 50 soorten in een opname. Bij de vochtindicatie volgens het ecotopensysteem ligt dit gemiddelde iets hoger: gemiddeld 18,3 soorten met indicatiewaarde, met een minimum van 3 soorten en een maximum van 52 soorten per opname.

Van alle 74 opnamen van Schaffers is het gemiddelde aantal soorten per opname 39,3. Hiervan kan 1 opname niet gebruikt worden voor de berekening van de correlatie tussen gemiddelde ecotopensysteem-indicatiewaarde en pH, omdat er volgens de gestelde criteria (paragraaf 2.3.4.1) geen soorten met indicatiewaarde aanwezig zijn. In de gebruikte opnamen zijn gemiddeld 12,5 soorten met ecotopensysteem-indicatiewaarde aanwezig met een minimum van 1 en een maximum van 33. Bij de berekeningen met Ellenberg R-getallen kunnen alle opnamen gebruikt worden. Het gemiddelde aantal soorten met indicatiewaarde per opname is hier veel hoger dan bij het ecotopensysteem: gemiddeld 23,0 soorten, met een minimum van 8 en een maximum van 39 soorten.

Formules waar naar verwezen wordt zijn te vinden in Bijlage 1 op pagina 73.

3.1.2 Vochtindicatie

3.1.2.1 Optimalisatie *BWW*-functie

In Bijlage 2 zijn de correlatiecoëfficiënten van alle waarden van a_0 , a_1 en a_2 weergegeven. De *BWW*-functie volgens formule (4) is het gevoeligst voor de waarde van de parameter a_0 . Voor de correlatie met GVG ligt bij zowel de Ellenberg als de ecotopensysteem-indicatiewaarden het optimum van a_0 in een vrij breed gebied (bij Ellenberg bij de waarden 0,189, 0,414 en 0,682 en bij het ecotopensysteem bij de waarden 0,189 en 0,414).

De resultaten zijn niet gevoelig voor variaties van a_1 en a_2 . Wat opvalt, is dat de waarde van a_0 effect heeft op de gevoeligheid voor de andere parameters. In het gebied waar de optimale waarde van a_0 ligt, is de spreiding van de correlatiecoëfficiënten als gevolg van de andere parameters klein, en komen er

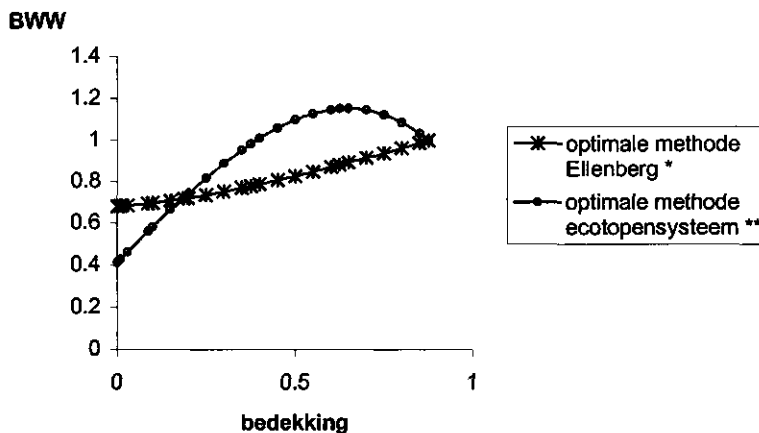
nauwelijks veel slechtere correlaties voor. De geschiktheid van de *BWW*-functie lijkt dus vooral bepaald te worden door de waarde van de parameter a_0 . Deze parameter bepaalt de grootte van het intercept, en daarmee grotendeels de verhouding tussen de hoogste en laagste *BWW* (ook de hoogte van de top in de *BWW*-functie bepaalt een gedeelte van de verhouding tussen hoogste en laagste gewicht).

De optimale methode voor Ellenberg levert bij het ecotopensysteem een net niet significant ($p = 5,1\%$) slechter resultaat dan de optimale methode voor het ecotopensysteem zelf. Andersom is, wanneer gewogen gemiddelde Ellenberg indicatiewaarden berekend worden, het verschil tussen de optimale methode voor Ellenberg en de optimale methode voor het ecotopensysteem totaal niet significant ($p = 48,1\%$). De optimale methode bepaald met het ecotopensysteem kan dus ook gebruikt worden voor Ellenberg F-getallen.

Voor het berekenen van een gemiddelde vochtindicatie, zowel met Ellenberg F-getallen als met het ecotopensysteem, is de *BWW*-functie met $a_0 = 0,4$; $a_1 = 1,5$ en $a_2 = 3,0$, verkregen door optimalisatie met het ecotopensysteem (Figuur 9), het beste. Ingevuld in formule (4) geeft dit:

$$BWW_i = 0,4 + 1,5 \left(\frac{B_i - B_{\min}}{B_{\max} - B_{\min}} \right) - 0,9 \left(\frac{B_i - B_{\min}}{B_{\max} - B_{\min}} \right)^3. \quad (18)$$

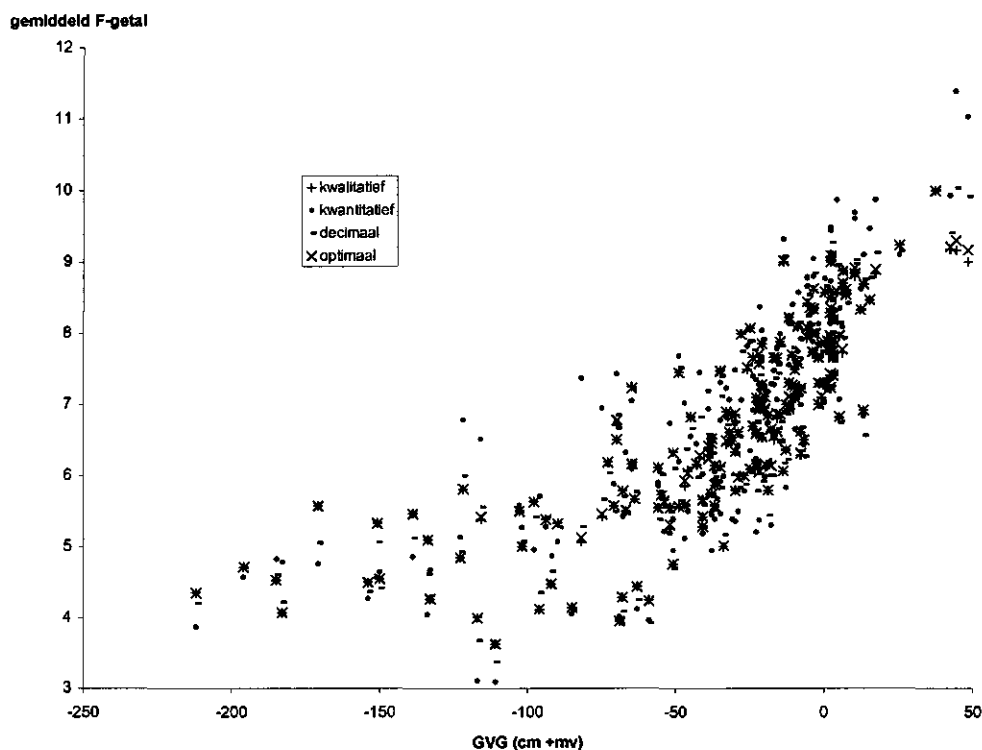
Deze methode geeft soorten met het de laagste bedekking ongeveer een drie keer zo klein gewicht als de soorten met het grootste gewicht (soorten die ongeveer $\frac{2}{3}$ van een opname bedekken (Braun-Blanquet code 4, zie Tabel 4). Dat soorten die $\frac{2}{3}$ van een opname bedekken een groter gewicht krijgen dan soorten die de opname volledig bedekken (Braun-Blanquet code 5 = B_{\max}) is waarschijnlijk het gevolg van toeval: de *BWW*-functie was immers niet erg gevoelig voor de waarden van de parameters a_1 en a_2 , die de ligging van deze top grotendeels bepalen.



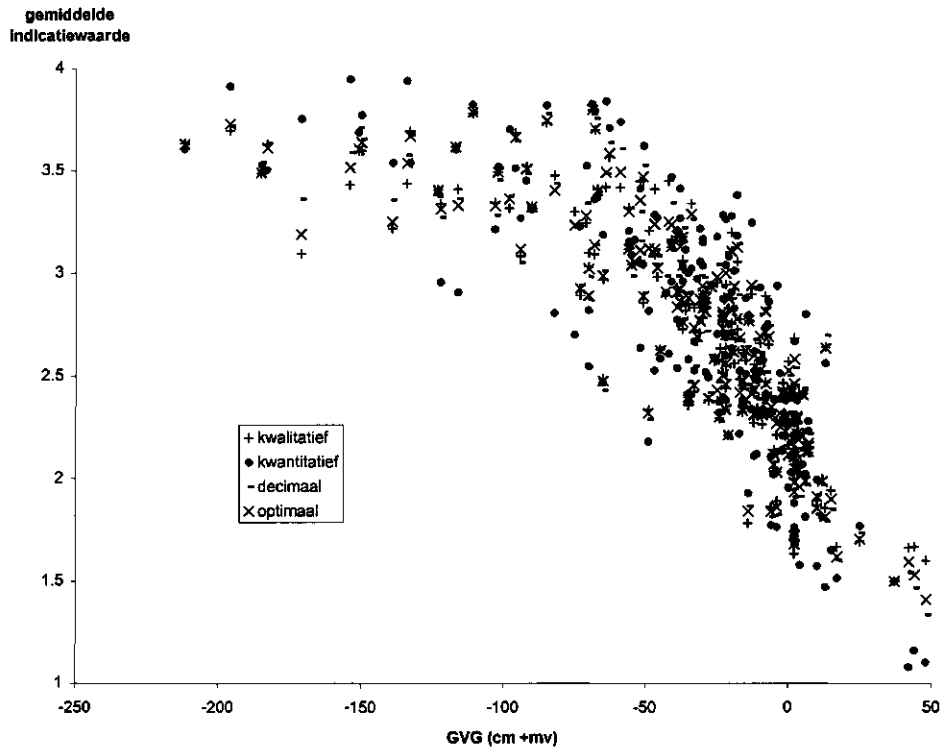
Figuur 9: Optimale bedekkingsweegmethoden voor de correlatie van de GVG met de gemiddelde indicatiewaarden volgens Ellenberg en het ecotopensysteem.*: $a_0 = 0,682$; $a_1 = -0,023$; $a_2 = 1,378$. **: $a_0 = 0,414$; $a_1 = 1,493$; $a_2 = 3,000$.

3.1.2.2 De beste methode

Het gebruik van de kwalitatieve methode leidt bij de vochtindicatie tot significant slechtere resultaten dan de optimale methode (Tabel 13 en Tabel 14 voor Ellenberg, Tabel 15 en Tabel 16 voor het ecotopensysteem). Uit dezelfde tabellen blijkt dat, wanneer er de keuze is tussen de drie onderzochte bestaande methoden, het niet uitmaakt of de kwalitatieve of decimale methode gebruikt wordt. In dat geval zou de kwalitatieve methode de voorkeur verdienen, omdat er dan geen bedekkingen geschat hoeven te worden. De kwantitatieve methode is voor beide indicatiesystemen wat de vochtindicatie betreft een slechte keus, hetgeen ook te zien is aan de grote spreiding in de gemiddelde indicatiewaarden volgens deze methode in Figuur 10 en Figuur 11. Bij de berekening met Ellenberg F-getallen verschilde de decimale methode niet significant van de optimale methode, bij de berekening met het ecotopensysteem wel. De optimale functie (18) is, ondanks één niet-significant verschil, het best. Deze functie blijkt in ieder geval nergens significant slechter te presteren dan andere methoden.



Figuur 10: Gemiddeld Ellenberg F-getal volgens vier BWW-methoden, uitgezet tegen de GVG (gegevens Runhaar).



Figuur 11: Gemiddelde ecotopensysteem-indicatiewaarde voor vocht berekend volgens vier BWW-methoden, uitgezet tegen de GVG (gegevens Runhaar).

Tabel 13: Spearman rangcorrelatie (r_s) tussen GVG en gemiddelde Ellenberg indicatiewaarde volgens aangegeven methode.

methode	r_s
kwalitatieve	0,868
kwantitatieve	0,832
decimale	0,863
optimale voor Ellenberg vocht †	0,872

†: $a_0 = 0,682$; $a_1 = -0,023$; $a_2 = 1,378$

Tabel 14: Overschrijdingskansen volgens de permutatietest voor het verschil tussen de correlatiecoëfficiënten van gemiddelde Ellenberg indicatiewaarden met GVG. Significante verschillen zijn dik gedrukt. Nulhypothese 'geen verschil' getest tegen alternatieve hypothese 'de optimale methode is beter' (*) of 'de methoden verschillen' (**).

methode	kwalitatieve	kwantitatieve	decimale	optimale voor Ellenberg vocht †
kwalitatieve	X	3,4%**	28,6%**	0,9%*
kwantitatieve		X	3,4%**	1,0%*
decimale			X	10,3%*

†: $a_0 = 0,682$; $a_1 = -0,023$; $a_2 = 1,378$

Tabel 15: Spearman rangcorrelatie (r_s) tussen GVG en gemiddelde ecotopensysteem-indicatiewaarde volgens aangegeven methode.

methode	r_s
kwalitatieve	-0,887
kwantitatieve	-0,833
decimale	-0,885
optimale ecotopensysteem vocht ‡	-0,895

‡: $a_0 = 0,414$; $a_1 = 1,493$; $a_2 = 3,000$

Tabel 16: Overschrijdingskansen voor de permutatietest voor het verschil tussen de correlatiecoëfficiënten van gemiddelde ecotopensysteem-indicatiewaarden met GVG. Significante verschillen zijn dik gedrukt. Nulhypothese 'geen verschil' getest tegen alternatieve hypothese 'de optimale methode is beter' (*) of 'de methoden verschillen' ().**

methode	kwalitatieve	kwantitatieve	decimale	optimale voor ecotopensysteem vocht ‡
kwalitatieve	X	0,8%**	73,2%**	2,0%*
kwantitatieve		X	0,2%**	0,0%*
decimale			X	0,7%*

‡: $a_0 = 0,414$; $a_1 = 1,493$; $a_2 = 3,000$

3.1.3 Zuurgraadindicatie

3.1.3.1 Optimalisatie *BWW*-functie

Bij de correlaties met pH ligt het optimum, zowel gebruik makend van Ellenberg als ecotopensysteem-indicatiewaarden, tussen 0,000 en 0,189 (Bijlage 2). De waarden van de parameters van de *BWW*-functies uit Tabel 6 blijken voor het vinden van de optimale functies bij de zuurgraadindicatie niet voldoende. Als aanvulling zijn met intervallen van 0,010 waarden van a_0 tussen 0,000 en 0,500 onderzocht (grafieken niet weergegeven in Bijlage 2).

De waarden van a_0 in de optimale *BWW*-functies voor zuurgraadindicatie waren erg laag: bij Ellenberg 0,020 (Tabel 17), bij het ecotopensysteem 0,050 (Tabel 19). Blijkbaar komen er veel toevallige, weinig indicatieve soorten voor, en zijn de soorten met de hoogste bedekkingen ook het meest indicatief. Helemaal niet meewegen van soorten met zeer lage bedekkingen ($a_0 = 0$) leidt weer tot erg slechte resultaten (Bijlage 2).

3.1.3.2 De beste methode

Wanneer gebruik wordt gemaakt van de Ellenberg R-getallen, is de kwantitatieve methode significant slechter dan de optimale (Tabel 18), bij het ecotopensysteem is dat het geval met de kwalitatieve methode, die daarnaast significant slechter is dan de decimale methode (Tabel 20). De decimale methode lijkt voor beide indicatiesystemen even goed te zijn als de optimale methoden. Dit alles maakt het

lastig om bij de zuurgraadindicatie een beste methode aan te wijzen. De kwantitatieve is bij Ellenberg niet goed, de kwalitatieve bij het ecotopensysteem. Verder is er geen methode die opvalt door significant hoge correlaties met pH.

Tabel 17: Spearman rangcorrelatie (r_s) tussen pH en gemiddelde Ellenberg indicatiewaarde volgens aangegeven methode.

methode	r_s
kwalitatieve	0,827
kwantitatieve	0,793
decimale	0,823
optimale voor Ellenberg pH ‡	0,853

‡: $a_0 = 0,020$; $a_1 = -0,023$; $a_2 = 2,364$

Tabel 18: Overschrijdingskansen voor de permutatietest voor het verschil tussen de correlatiecoëfficiënten van gemiddelde Ellenberg indicatiewaarden met pH. Significante verschillen zijn dik gedrukt. Nulhypothese 'geen verschil' getest tegen alternatieve hypothese 'de optimale methode is beter' (*) of 'de methoden verschillen' ().**

methode	kwalitatieve	kwantitatieve	decimale	optimale voor Ellenberg pH ‡
kwalitatieve	X	36,0%**	64,4%**	8,7%*
kwantitatieve		X	35,2%**	3,2%*
decimale			X	5,6%*

‡: $a_0 = 0,020$; $a_1 = -0,023$; $a_2 = 2,364$

Tabel 19: Spearman rangcorrelatie (r_s) tussen pH en gemiddelde ecotopensysteem-indicatiewaarde volgens aangegeven methode.

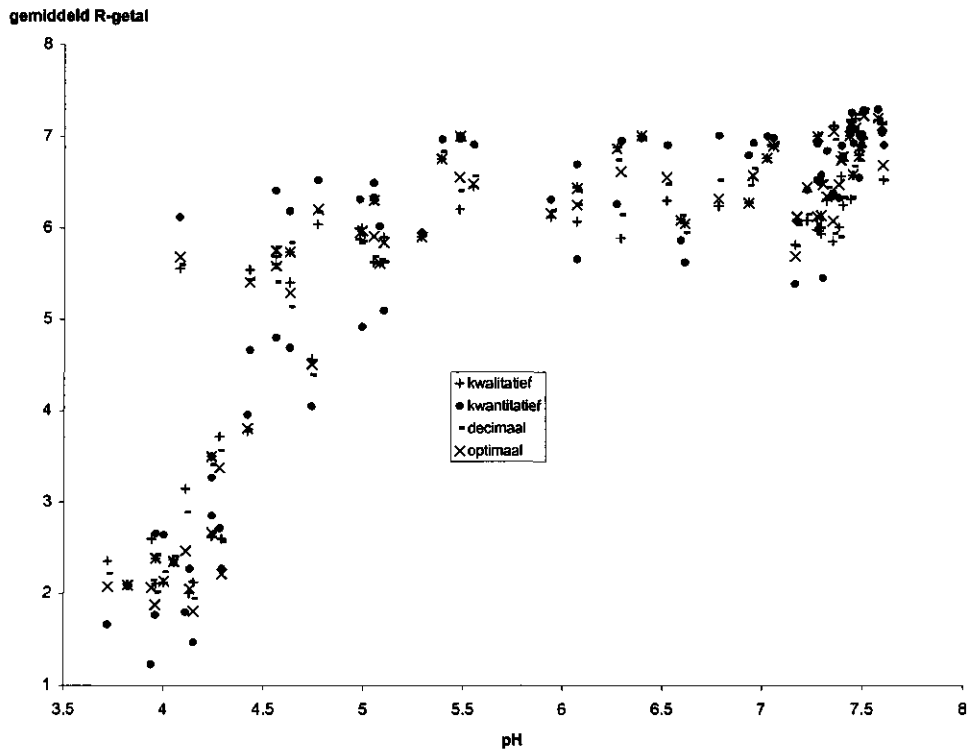
methode	r_s
kwalitatieve	0,834
kwantitatieve	0,847
decimale	0,853
optimale voor ecotopensysteem pH ‡	0,869

‡: $a_0 = 0,050$; $a_1 = 0,187$; $a_2 = 1,000$

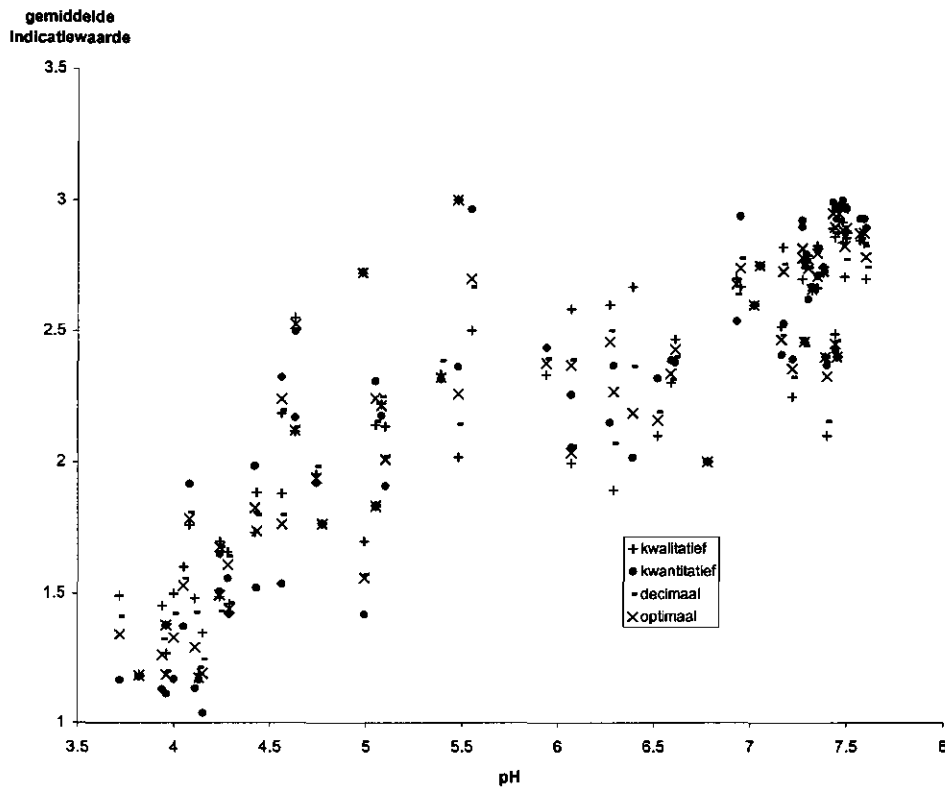
Tabel 20: Overschrijdingskansen voor de permutatietest voor het verschil tussen de correlatiecoëfficiënten van gemiddelde ecotopensysteem-indicatiewaarden met pH. Significante verschillen zijn dik gedrukt. Nulhypothese 'geen verschil' getest tegen alternatieve hypothese 'de optimale methode is beter' (*) of 'de methoden verschillen' ().**

methode	kwalitatieve	kwantitatieve	decimale	optimale voor ecotopensysteem pH ‡
kwalitatieve	X	49,8%**	3,2%**	0,2%*
kwantitatieve		X	76,8%**	14,0%*
decimale			X	7,4%*

‡: $a_0 = 0,050$; $a_1 = 0,187$; $a_2 = 1,000$



Figuur 12: Gemiddeld Ellenberg R-getal berekend met vier BWV-methoden, uitgezet tegen de pH (gegevens Schaffers).



Figuur 13: Gemiddelde ecotoposysteem-indicatievalue voor zuurgraad berekend volgens vier BWV-methoden, uitgezet tegen de pH (gegevens Schaffers).

3.2 Ellenberg versus ecotopensysteem

In Tabel 21 en Tabel 22 zijn de verschillen tussen de rangcorrelatiecoëfficiënten van Ellenberg en het ecotopensysteem per weegmethode weergegeven. De resultaten zijn vrij eenduidig: bij de correlatie met GVG geeft het ecotopensysteem de hoogste correlaties (alleen bij de kwantitatieve methode niet significant, maar de kwantitatieve methode doet het meestal vrij slecht en is dus van weinig belang), bij de correlatie met pH is geen enkel verschil significant. Op het oog lijkt de relatie tussen indicatiewaarde en pH bij hogere pH's bij het ecotopensysteem (Figuur 13) beter dan bij de Ellenberg R-getallen (Figuur 12), maar dit uit zich niet in een significant hogere correlatiecoëfficiënt, waarschijnlijk doordat de spreiding in gemiddelde indicatiewaarden bij het ecotopensysteem groter is dan bij Ellenberg. Voor vochtindicatie kan dus beter het ecotopensysteem gebruikt worden, voor de zuurgraadindicatie lijkt het niet uit te maken.

Tabel 21: Vergelijking tussen de correlatie van gemiddelde Ellenberg indicatiewaarden en gemiddelde ecotopensysteem-indicatiewaarden met de GVG. r_s is geschat volgens de bootstrapmethode. De (tweezijdige) overschrijdingskansen zijn berekend met de permutatietest. Significante verschillen zijn vet gedrukt. Getest is de nulhypothese 'geen verschil' tegen de alternatieve hypothese 'wel verschil'.

methode	r_s Ellenberg	r_s ecotopensysteem	overschrijdingskans
kwalitatieve	0,868	-0,887	3,6%
kwantitatieve	0,832	-0,833	91,4%
decimale	0,863	-0,885	2,0%
optimale voor Ellenberg vocht ‡	0,872	-0,890	3,2%
optimale voor ecotopensysteem vocht †	0,871	-0,895	0,2%

‡: $a_0 = 0,682$; $a_1 = -0,023$; $a_2 = 1,378$

†: $a_0 = 0,414$; $a_1 = 1,493$; $a_2 = 3,000$

Tabel 22: Vergelijking tussen de correlatie van gemiddelde Ellenberg indicatiewaarden en gemiddelde ecotopensysteem-indicatiewaarden met de pH. r_s is geschat volgens de bootstrapmethode. De (tweezijdige) overschrijdingskansen zijn berekend met de permutatietest. Significante verschillen zijn vet gedrukt. Getest is de nulhypothese 'geen verschil' tegen de alternatieve hypothese 'wel verschil'.

methode	r_s Ellenberg	r_s ecotopensysteem	overschrijdingskans
kwalitatieve	0,827	0,834	84,4%
kwantitatieve	0,793	0,847	25,2%
decimale	0,823	0,853	44,8%
optimale voor Ellenberg pH ‡	0,853	0,863	54,4%
optimale voor ecotopensysteem pH †	0,831	0,869	5,4%

‡: $a_0 = 0,020$; $a_1 = -0,023$; $a_2 = 2,364$

†: $a_0 = 0,050$; $a_1 = 0,187$; $a_2 = 1,000$

3.3 Soortspecifieke bedekkingsweegwaarden

In Tabel 24, Tabel 26, en Tabel 28 is te zien dat de optimale soortspecifieke bedekkingsweegwaardefunctie in de meeste gevallen geen erg groot, positief effect heeft op de correlaties van gemiddelde indicatiewaarde met het milieu. Wanneer Ellenberg F-getallen gebruikt worden (resultaten niet weergegeven), is dit het duidelijkst: de optimale waarde van de parameter a_0 is 1,0, hetgeen overeen komt met het geven van een gelijk gewicht aan alle soorten, ongeacht hun soortspecifieke bedekkingseigenschappen (de kwalitatieve methode dus). Witte (ongepubliceerd) kwam, gebruik makend van dezelfde gegevens, al eerder tot deze conclusie.

Of nu de gemiddelde bedekking of de gemiddeld maximale bedekking gebruikt wordt in de soortspecifieke *BWW*-functie, maakt niets uit. Weliswaar verschillen de correlatiecoëfficiënten en de optimale waarden van de parameters a_0 en a_1 (Tabel 23, Tabel 25 en Tabel 27), maar het effect is steeds niet significant (Tabel 24, Tabel 26 en Tabel 28).

Alleen bij de zuurgraadindicatie met Ellenberg R-getallen is de soortspecifieke *BWW*-functie in enkele gevallen significant beter dan een standaard *BWW*-methode (Tabel 24). De optimale waarde van a_0 (0,9) zorgt er echter voor, dat de verschillen tussen de gewichten van de afzonderlijke soorten in deze optimale methode klein zijn. Ik concludeer hier voorlopig uit dat ook bij Ellenberg R-getallen geen sterk positief effect van soortspecifieke bedekkingen aanwezig is. Het significante verschil in sommige gevallen zal eerder veroorzaakt worden door toevalligheden in de gegevens waaruit de soortspecifieke bedekkingen bepaald zijn.

Tabel 23: Optimale soortspecifieke bedekkingsweegwaardefuncties voor Ellenberg R-getallen bij verschillende maten voor soortspecifieke bedekking. GMB = gemiddeld maximale bedekking, berekend over het tussen haakjes aangegeven percentage hoogste bedekkingen van elke soort. Spearman rangcorrelatie (r_s) tussen pH en gemiddelde Ellenberg indicatiewaarde is geschat met de bootstrap.

soortspecifieke maat	a_0	a_1	r_s
gemiddelde bedekking	0,9	1,96	0,857
GMB (25%)	0,9	3,24	0,871
GMB (10%)	0,9	4,00	0,861

Tabel 24: Tweezijdige overschrijdingskansen voor de permutatietest voor het verschil tussen de correlatiecoëfficiënten van gemiddelde Ellenberg R-getallen met pH. Significante verschillen zijn dik gedrukt. Getest is de nulhypothese 'geen verschil' tegen de alternatieve hypothese 'wel verschil'. Kwalitatieve, decimale en optimale methode zijn bedekkingsweegwaardefuncties waarbij geen rekening gehouden is met de soortspecifieke bedekking.

	gemiddelde bedekking	GMB (25%)	GMB (10%)	kwalitatieve methode	decimale methode	optimale methode Ellenberg pH zonder verfijning ‡
gemiddelde bedekking	X	9,4%	73,8%	8,2%	3,4%	17,0%
GMB (25%)		X	23,4%	3,2%	1,8%	7,4%
GMB (10%)			X	13,2%	8,0%	20,4%

‡ $a_0 = 0,189$; $a_1 = -0,630$; $a_2 = 1,378$

Tabel 25: Optimale soortspecifieke bedekkingsweegwaardefuncties voor ecotopensysteem-vochtindicatie bij verschillende maten voor soortspecifieke bedekking. GMB = gemiddeld maximale bedekking, berekend over het tussen haakjes aangegeven percentage hoogste bedekkingen van elke soort. Spearman rangcorrelatie (r_s) tussen GVG en gemiddelde ecotopensysteem-indicatie waarde is geschat met de bootstrap.

soortspecifieke maat	a_0	a_1	r_s
gemiddelde bedekking	0,8	0,64	-0,890
GMB (25%)	0,8	0,16	-0,890
GMB (10%)	0,8	0,16	-0,890

Tabel 26: Tweezijdige overschrijdingskansen voor de permutatietest voor het verschil tussen de correlatiecoëfficiënten van gemiddelde ecotopensysteem-vochtindicatie waarden met GVG. Significante verschillen zijn dik gedrukt. Getest is de nulhypothese 'geen verschil' tegen de alternatieve hypothese 'wel verschil'. Kwalitatieve, decimale en optimale methode zijn bedekkingsweegwaardefuncties waarbij geen rekening gehouden is met de soortspecifieke bedekking.

	gemiddelde bedekking	GMB (25%)	GMB (10%)	kwalitatieve methode	decimale methode	optimale methode ecotopensysteem vocht ‡
gemiddelde bedekking	X	12,2%	23,8%	15,8%	28,6%	11,2%
GMB (25%)		X	6,3%	71,6%	62,4%	4,0%
GMB (10%)			X	46,1%	50,0%	6,0%

‡: $a_0 = 0,414$; $a_1 = 1,493$; $a_2 = 3,000$

Tabel 27: Optimale soortspecifieke bedekkingsweegwaardefuncties voor ecotopensysteem-zuurgraadindicatie bij verschillende maten voor soortspecifieke bedekking. GMB = gemiddeld maximale bedekking, berekend over het tussen haakjes aangegeven percentage hoogste bedekkingen van elke soort. Spearman rangcorrelatie (r_s) tussen pH en gemiddelde ecotopensysteem-indicatiewaarde is geschat met de bootstrap.

soortspecifieke maat	a_0	a_1	r_s
gemiddelde bedekking	0,4	0,64	0,841
GMB (25%)	0,3	1,00	0,843
GMB (10%)	0,1	1,00	0,852

Tabel 28: Tweezijdige overschrijdingskansen voor de permutatietest voor het verschil tussen de correlatiecoëfficiënten van gemiddelde ecotopensysteem-zuurgraadindicatiewaarden met pH. Significante verschillen zijn dik gedrukt. Getest is de nulhypothese 'geen verschil' tegen de alternatieve hypothese 'wel verschil'. Kwalitatieve, decimale en optimale methode zijn bedekkingsweegwaardefuncties waarbij geen rekening gehouden is met de soortspecifieke bedekking.

	gemid- delde bedekking	GMB (25%)	GMB (10%)	kwali- tatieve methode	decimale methode	optimale methode ecotopensysteem pH zonder verfijning ‡
gemiddelde bedekking	X	96,0 %	56,0%	52,8%	20,0%	29,2%
GMB (25%)		X	43,8%	63,6%	42,2%	42,8%
GMB (10%)			X	43,6%	77,8%	72,2%

‡: $a_0 = 0,189$; $a_1 = -0,630$; $a_2 = 1,378$

3.4 Indicatieweegwaarden (Ellenberg)

Indicatieweegwaarden bepalen het gewicht van soorten met een 'extreme' indicatiewaarde ten opzichte van soorten met een meer 'gematigde' indicatiewaarde, afhankelijk van het aantal indifferenten in de opname. Bij de vochtindicatie neemt de correlatiecoëfficiënt door het gebruik van *IWW*'s volgens de formules (7) en (8) bijna overal significant toe, behalve wanneer de decimale methode gebruikt wordt (Tabel 29). Bij de zuurgraadindicatie is geen significant verschil te ontdekken (Tabel 30). Bij de zuurgraadindicatie is niet meer gezocht naar een optimale combinatie van *BWW*- en *IWW*-functie voor a_0 tussen 0,000 en 0,189, vanwege het ontbreken van enig significant effect in Tabel 30. Figuur 14 laat zien waaruit de invloed van de *IWW* op de gemiddelde vochtindicatiewaarden bestaat: vooral opnamen met een grondwaterstand boven het maaiveld krijgen een hogere gemiddelde indicatiewaarde. De opnamen met een diepe grondwaterstand krijgen iets lagere gemiddelde indicatiewaarden, maar extreem droge opnamen, waar het effect het sterkst zou moeten zijn, ontbreken.

De optimale waarden voor zowel c_1 als c_2 uit formule (7) zijn bij de vochtindicatie hoger dan bij de zuurgraadindicatie. De lagere waarden voor c_1 bij de zuurgraadindicatie zouden samen kunnen hangen met het ontbreken van een

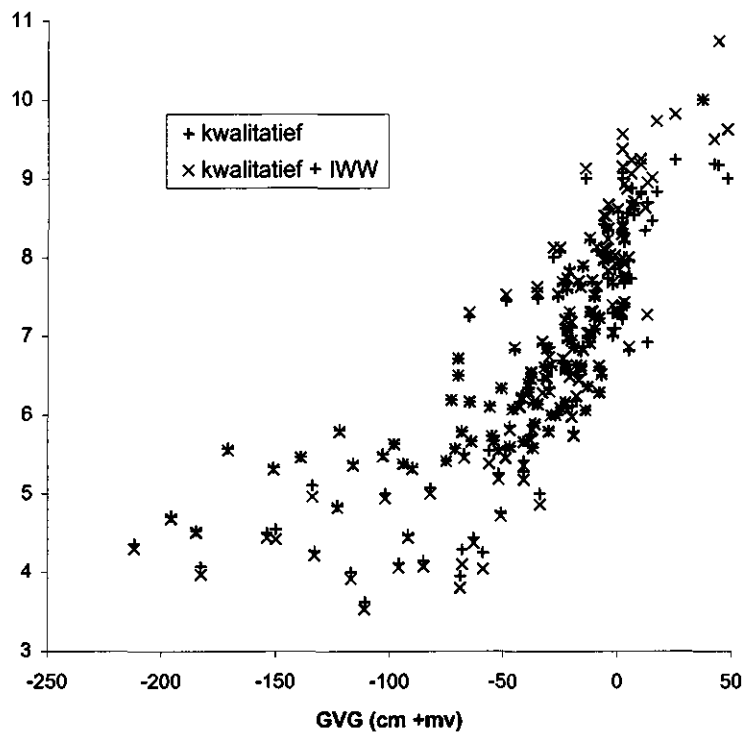
significant effect *überhaupt*. In de grafieken in Bijlage 3 is te zien dat het effect van zowel c_1 als c_2 bij de vochtindicatie groter is dan bij de zuurgraadindicatie. Aangezien het verlenen van gewichten aan indicatiewaarden overeen komt met het 'uitrekken' van de schaal bij de extreme indicatiewaarden (het vergroten van de afstand tussen de extreme indicatiewaarden) kan dit een aanwijzing zijn dat de F-getallen van Ellenberg de soorten niet goed van elkaar onderscheiden. Er zijn te veel klassen voor soorten met een ongeveer gemiddeld voorkomen, of te weinig voor soorten met een extreem optimum. De R-getallen zouden dan beter aan de soorten toebedeeld zijn dan met de F-getallen het geval is, vanwege het ontbreken van een significant effect van de indicatieweegwaarden bij de R-getallen.

Bij de vochtindicatie is de optimale waarde van de parameter c_2 hoog, en daarmee is het aantal indifferenten belangrijk. Slechts een lage bedekking van indifferenten in de opname zorgt al voor een grote afname in het gewicht van de extreme indicatiewaarden ten opzichte van de situatie zonder indifferenten. Wanneer bij de kwalitatieve en optimale *BWW*-methoden de optimale c_1 en c_2 gekozen worden, is de correlatie significant beter dan wanneer voor c_2 de waarde 0 gekozen wordt (respectievelijk $p = 0,0\%$ en $p = 0,5\%$). Dit blijkt ook uit de eerste grafiek in Bijlage 3: hoe hoger de waarde van c_1 , des te meer invloed heeft c_2 . Blijkbaar is het verlenen van een groot gewicht aan extreme indicatiewaarden alleen maar nuttig als er weinig tot geen indifferenten in de opname aanwezig zijn. Figuur 15 geeft een verklaring: alle opnamen zonder indifferenten hebben in het voorjaar het grondwater boven het maaiveld. De theorie dat indifferenten extreme omstandigheden mijden lijkt voor erg natte omstandigheden in ieder geval te kloppen. Erg droge opnamen ontbreken in de gegevens.

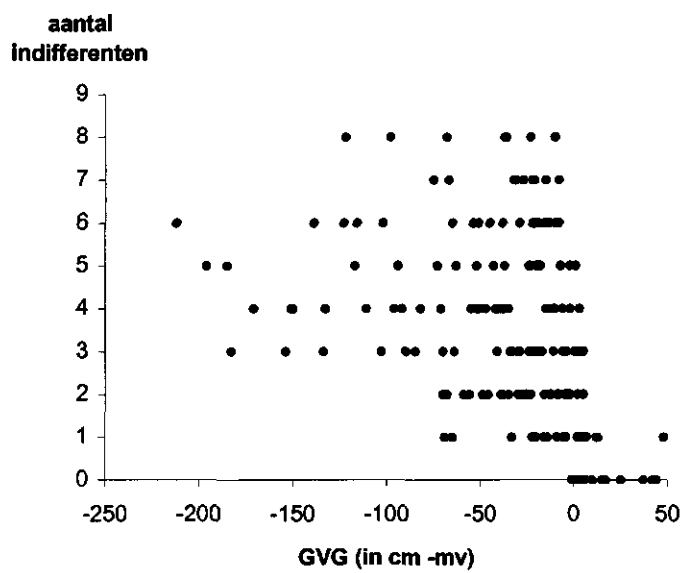
In de gegevens van Schaffers, die gebruikt zijn voor het berekenen van de correlaties met pH, zijn geen opnamen zonder indifferenten aanwezig. Wel is in Figuur 16 te zien dat bij hoge of lage pH veel minder indifferenten voorkomen (wanneer het aandeel indifferenten van het totaal aantal soorten bekeken wordt, ontstaat hetzelfde beeld). Blijkbaar voegen de indifferente soorten hier geen informatie toe.

Twee onderdelen van de *IWW*-formule (7) zijn redelijk arbitrair, namelijk het kiezen van de middelste indicatiewaarde als indicatiewaarde die het laagste gewicht krijgt, en het kwadrateren van de afstand tot die middelste indicatiewaarde. Bij Durwen (1982 pag. 36; ook te vinden in Ellenberg 1992 pag. 44) is te zien dat ongeveer 30% van alle soorten een R-getal lager of gelijk aan 5 heeft, en meer dan 60 % van de soorten een F-getal lager dan de centrale waarde 6,5 voor vocht heeft. Als meest gematigde indicatiewaarde moet bij de R-getallen dus misschien een hogere waarde gekozen worden, en bij de F-getallen een lagere. Of daarnaast het kwadraat van de afstand tot de middelste indicatiewaarde beter is dan een andere transformatie, is nauwelijks te voorspellen op basis van de literatuur en kennis over indicatiewaarden.

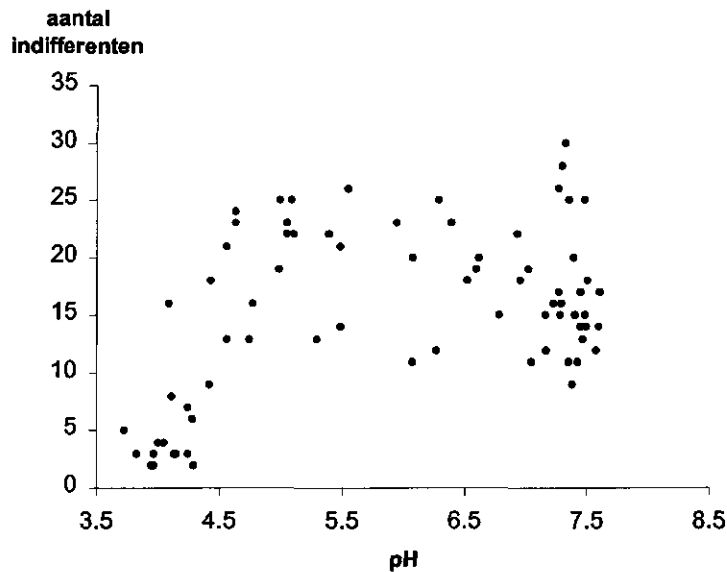
vochtindicatie



Figuur 14: Relatie tussen GVG en gemiddelde indicatiewaarden wanneer deze volgens de kwalitatieve methode met en zonder indicatieweegwaarden (IWW) berekend worden.



Figuur 15: Relatie tussen GVG en het aantal indifferenten in de gegevens van Runhaar.



Figuur 16: De relatie tussen het aantal indifferenten en de pH in de gegevens van Schaffers.

Tabel 29: De optimale waarden van de parameters in de indicatieweegwaardefunctie, per bedekkingsweegwaardefunctie. De Spearman rangcorrelatie (r_s) van gemiddelde Ellenberg indicatiewaarden met de GVG is geschat volgens de bootstrap-methode. De (eenzijdige) overschrijdingskansen zijn berekend met de permutatietest. Significante verschillen zijn vet gedrukt. Getest is de nulhypothese 'geen verschil' tegen de alternatieve hypothese 'de IWW-methode is beter'.

bedekkingsweegmethode	c_1	c_2	r_s	overschrijdingskans
kwalitatieve	38,037	28,000	0,875	2,8%
kwantitatieve	50,296	5,630	0,840	1,1%
decimale	65,000	19,963	0,869	10,0%
optimale voor Ellenberg IWW vocht ‡	38,037	50,296	0,878	4,4%

‡: $a_0 = 0,414$; $a_1 = 0,630$; $a_2 = 1,000$

Tabel 30: De optimale waarden van de parameters in de indicatieweegwaardefunctie, per bedekkingsweegwaardefunctie. De Spearman rangcorrelatie (r_s) van gemiddelde Ellenberg indicatiewaarden met de pH is geschat volgens de bootstrap-methode. De (eenzijdige) overschrijdingskansen zijn berekend met de permutatietest. Significante verschillen zijn vet gedrukt. Getest is de nulhypothese 'geen verschil' tegen de alternatieve hypothese 'de IWW-methode is beter'.

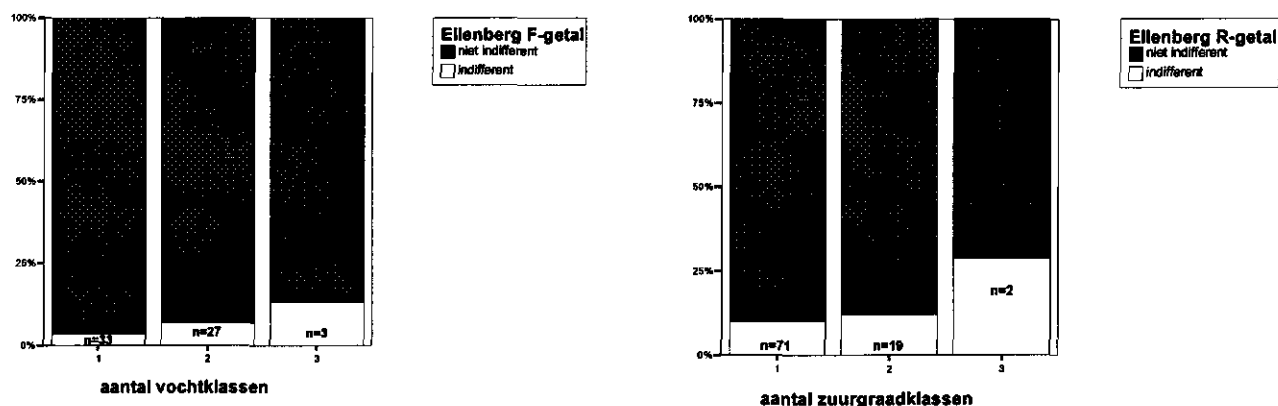
bedekkingsweegmethode	c_1	c_2	r_s	overschrijdingskans
kwalitatieve	28,000	0,963	0,839	29,0%
kwantitatieve	1,296	19,963	0,789	100,0%
decimale	19,963	2,000	0,832	25,6%
optimale voor Ellenberg IWW pH ‡	13,704	1,037	0,847	30,1%

‡: $a_0 = 0,189$; $a_1 = 1,493$; $a_2 = 0,682$

3.5 Soortweegwaarden (ecotopensysteem)

In de resultaten in Tabel 31 en Tabel 32 valt op dat het toepassen van soortweegwaarden (gebaseerd op de breedte van de ecologische amplitude zoals die beschreven wordt door het ecotopensysteem, formules (12) en (13)) nergens tot

significante verbetering van de correlatiecoëfficiënten leidt. Bij een vergelijking van de 'amplitude-maten' van Ellenberg (indifferente soorten hebben een erg brede amplitude) en het ecotopensysteem is te zien dat deze niet heel goed overeen komen (Figuur 17). Er zijn veel soorten die door Ellenberg als 'indifferent' ingedeeld zijn, maar door het ecotopensysteem slechts in één klasse. Waarschijnlijk gedraagt een gedeelte van de soorten die door Ellenberg van een indicatiewaarde zijn voorzien zich in Nederland anders dan in Midden-Europa (Ellenberg 1992). Misschien is de indeling van het ecotopensysteem te arbitrair, waardoor het effect van Figuur 6, waarbij twee soorten met gelijke amplitude in een verschillend aantal ecologische groepen ingedeeld worden, vaker optreedt dan verwacht. Het verschil tussen soorten die in één of twee groepen ingedeeld zijn wordt zo nietszeggend. De soorten die in meerdere groepen met betrekking tot één milieufactoor voor kunnen komen, worden in veel toepassingen van het ecotopensysteem niet meegenomen, in Runhaar et al. (1997) bijvoorbeeld omdat de resultaten beter werden wanneer deze soorten weggelaten werden.



Figuur 17: Aantal en percentage indifferente soorten bij Ellenberg-getallen per indeling in 1, 2 of 3 groepen van het ecotopensysteem. De staven in het diagram zijn geschaald naar 100%, waarbij van elke groep het aantal soorten (n) aangegeven is. a) (links) Ellenberg F-getallen en vochtgroepen ecotopensysteem. b) (rechts) Ellenberg R-getallen en zuurgraadgroepen ecotopensysteem. Soorten zonder indeling naar zuurgraad in het ecotopensysteem zijn niet weergegeven.

Tabel 31: De optimale waarden van de parameter d in de soortweegwaardefunctie, per bedekkingsweegwaardefunctie. De Spearman rangcorrelatie (r_s) van gemiddelde ecotopensysteem-indicatiewaarden met de GVG is geschat volgens de bootstrap-methode. De (eenzijdige) overschrijdingskansen zijn berekend met permutatietesten. Significante verschillen zijn vet gedrukt. Getest is de nulhypothese 'geen verschil' tegen de alternatieve hypothese 'de SWW-methode is beter'.

bedekkingsweegmethode	d	r_s	overschrijdingskans
kwalitatieve	1,088	-0,889	30,2%
kwantitatieve	3,370	-0,845	11,6%
decimale	1,011	-0,886	44,2%
optimale ecotopensysteem SWW vocht ‡	0,704	-0,895	29,7%

‡: $a_0 = 0,189$; $a_1 = -0,023$; $a_2 = 0,682$

Tabel 32: De optimale waarden van de parameter d in de soortweegwaardefunctie, per bedekkingsweegwaardefunctie. De Spearman rangcorrelatie (r_s) van gemiddelde ecotopensysteem-indicatiewaarden met de pH is geschat volgens de bootstrap-methode. De (eenzijdige) overschrijdingskansen zijn berekend met permutatietesten. Significante verschillen zijn vet gedrukt. Getest is de nulhypothese 'geen verschil' tegen de alternatieve hypothese 'de SWW-methode is beter'.

bedekkingsweegmethode	d	r_s	overschrijdingskans
kwalitatieve	0,912	0,838	19,0%
kwantitatieve	4,764	0,836	38,8%
decimale	0,912	0,857	24,1%
optimale ecotopensysteem SWW pH ‡	0,704	0,863	13,4%

‡: $a_0 = 0,000$; $a_1 = 0,630$; $a_2 = 0,414$

3.6 Mediaan

Wanneer de mediane Ellenberg indicatiewaarden per opname berekend worden met de formule voor de mediaan (15), levert dit een r_s van 0,850 op voor de correlatie met GVG, en een r_s van 0,783 voor de correlatie met pH. Deze correlaties van de mediane indicatiewaarden met het milieu zijn daarmee lager dan de correlaties van de ongewogen gemiddelde indicatiewaarden met het milieu: 0,868 voor vochtindicatie met GVG, en 0,827 voor zuurgraadindicatie met pH. Voor de vochtindicatie is de correlatiecoëfficiënt wanneer de mediaan gebruikt wordt significant lager dan wanneer de indicatie met de kwalitatieve methode wordt berekend ($p = 1,5\%$). Voor de zuurgraadindicatie is het verschil niet significant ($p = 13,0\%$). De resultaten van deze berekeningswijze voor de mediaan (formule (15)) verschillen niet significant van de standaardmethode, waarin de mediaan eenvoudig de middelste waarneming is ($p = 89,8\%$ voor pH, $p = 64,4\%$ voor GVG).

Hoewel het berekenen van de mediaan een statistisch correcte methode is en het berekenen van het gemiddelde niet, komen de gemiddelde indicatiewaarden meer overeen met gemeten milieufactoren dan de mediane indicatiewaarden. Ter Braak & Barendregt (1986) geven in hun theoretische uiteenzetting de voorwaarden waaraan het berekenen van gewogen gemiddelden moet voldoen, en de methoden om het gebruik van gewogen gemiddelden aanvaardbaar te maken. De realiteit lijkt echter niet in overeenstemming met de door hen gevonden voorwaarden (voor deze voorwaarden, zie Ter Braak & Barendregt 1986).

4 Belangrijkste conclusies en discussie

4.1 Enkele opmerkingen over statistiek

Het beeld van de natuur zoals wij dat voorgeschoteld krijgen is bepaald door een groot aantal, op complexe wijze met elkaar samenhangende factoren. Daarbij is de respons van soorten op al deze factoren ook nog eens stochastisch variabel. De invloed van de factoren die bijdragen aan het tot stand komen van de manier waarop ecosystemen in elkaar zitten zijn daarbij vrijwel nooit geheel duidelijk. Eén manier waarop het mogelijk is complexe ecosystemen te onderzoeken is door experimenten uit te voeren onder gecontroleerde omstandigheden. De onderzoeker bepaalt dan zelf welke factoren hij laat variëren tussen de verschillende behandelingen, en welke factoren constant blijven. Dit is echter voor veel ecologisch onderzoek eenvoudigweg onmogelijk, bijvoorbeeld vanwege hoge kosten of lange tijd en grote oppervlakten die de onderzoeker nodig heeft, of omdat het controleren van de omstandigheden onmogelijk is. Als gecontroleerde experimenten niet mogelijk zijn, is de onderzoeker genoodzaakt zijn toevlucht te nemen tot de statistische verwerking van waarnemingen. Met adequate statistiek kan (wanneer de juiste methoden toegepast worden) een duidelijk, te begrijpen beeld gegeven worden, doordat de werkelijkheid beschreven wordt in termen van gemiddelden, medianen, correlatiecoëfficiënten, regressielijnen, standaardfouten en betrouwbaarheidsintervallen. Meer in het algemeen is het met juist gebruik van statistiek mogelijk het effect van toeval te onderscheiden van de effecten van de factoren waarin men geïnteresseerd is.

Het gebruik van statistiek in de ecologie is echter niet eenvoudig. Veelal worden in het ecologisch onderzoek statistische methoden gebruikt die vereisen dat de populatie waaruit de waarnemingen genomen zijn normaal verdeeld is. Meestal is er echter weinig bekend over de verdeling van de statistische populatie. Wanneer er wel iets bekend is over de specifieke verdeling, kan een transformatie uitgevoerd worden om de waarnemingen meer op een normale verdeling te laten lijken. Inherent aan toetsen die uitgaan van normale verdelingen is echter dat de uitspraken van de toetsen beïnvloed worden door transformaties. Hiermee wordt de uitspraak van de toets afhankelijk van de keuze van de onderzoeker voor een transformatie. En vaak is het zelfs met behulp van een transformatie niet mogelijk om de gegevens bij benadering normaal verdeeld te krijgen, zodat de uitkomsten van de toets altijd in meer of mindere mate fout zijn.

Een ander probleem bij het gebruik van klassieke statistiek in de ecologie vormt de voorwaarde dat de steekproeven aselekt en willekeurig moeten zijn. Vaak is het niet mogelijk om een willekeurige, aselechte steekproef te nemen (Witte & Runhaar 2000, Edgington 1995). Aselechte steekproeven zijn wel vereist voor het gebruik van statistische tabellen bij klassieke statistische toetsen, zoals de t- en F-tabellen, waarin de kritieke waarden voor de onbetrouwbaarheidsdrempels gebaseerd zijn op willekeurige steekproeven uit een theoretische (meestal normale) verdeling. Het

gebruik van dergelijke statistische tabellen leidt, als er niet aan de voorwaarde van een aselechte, willekeurige steekproef is voldaan, tot verkeerde conclusies. Ook het gebruik van tabellen met kritieke waarden van de zogenaamde verdelingsvrije methoden (die al een groot voordeel hebben boven de klassieke statistiek omdat ze niet uitgaan van een specifieke verdeling) vereist nog steeds aselechte steekproeven.

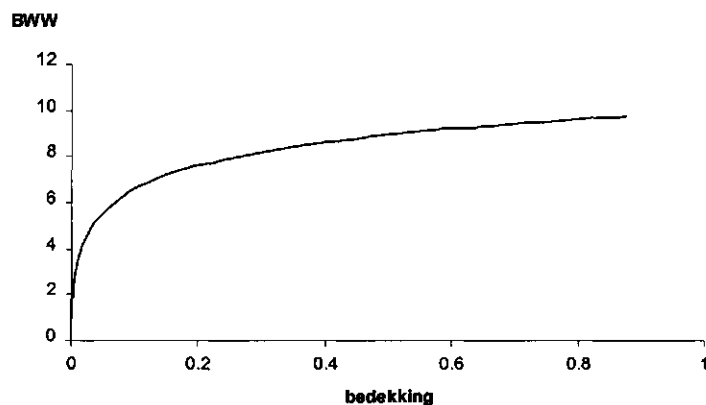
Wanneer geen aselechte steekproeven genomen zijn, kan er geen uitspraak gedaan worden over de populatie waaruit de steekproeven genomen zijn. Het blijft echter wel wenselijk om te kunnen bepalen of de resultaten van het onderzoek statistisch significant zijn binnen de steekproef. Omdat klassieke tabellen met kritieke waarden hiervoor niet gebruikt kunnen worden, moeten de kritieke waarden en overschrijdingskansen uit de gegevens worden berekend. Permutatietesten en andere *resampling* technieken zoals de bootstrap zijn hierbij de juiste methoden (Edgington 1995). Deze methoden worden in de ecologie echter weinig toegepast, terwijl het gebruik van methoden die uitgaan van aselechte steekproeven (en die dus altijd min of meer verkeerde uitkomsten geven) ten onrechte algemeen geaccepteerd lijkt.

4.2 Het berekenen van gemiddelde indicatiewaarden

In deze studie is aangetoond dat er, tenminste voor de hier gebruikte vegetatieopnamen van Runhaar, betere methoden zijn dan de standaard methoden om de gemiddelde vochtindicatiewaarde van een vegetatieopname te berekenen. Voor de zuurgraadindicatie zijn geen duidelijk betere methoden gevonden. Om de vochtindicatie van een vegetatieopname te berekenen, is het ecotopensysteem beter dan de indicatiewaarden van Ellenberg. Bij beide indicatiesystemen is er een bedekkingsweegwaardemethode te vinden (formule (18)) die betere resultaten oplevert dan de kwalitatieve en kwantitatieve methode, en die bij het ecotopensysteem ook beter is dan de decimale methode. Deze optimale methode wordt gekenmerkt door een vrij klein verschil tussen maximaal en minimaal gewicht. In dit opzicht verschilt de optimale methode van de methode die door het CML voorgesteld is voor het ecotopensysteem ($BWW_i = {}^2\log[1000*B_i + 1]$, bijlage 5.2 in Groen et al. 1993, Figuur 18). De optimale methode uit deze studie komt meer overeen met de methoden die door de provincies Zuid-Holland (Clausman & Van Wijngaarden 1984) en Gelderland (Hertog & Rijken 1992) voor natuurwaardering gebruikt worden, omdat in die methoden ook gebruik wordt gemaakt van een korte *BWW*-schaal (resp. van ¼ en ½ tot 1).

Kunzmann (1989) laat enkele voorbeelden van plantensoorten zien waarbij er een verschil is tussen presentieoptimum (gevormd door de omstandigheden waarbij een soort het vaakst aangetroffen wordt) en bedekkingsoptimum (de omstandigheden waarbij de soort de hoogste bedekkingen bereikt). Ook van soorten in Nederland is dergelijk gedrag bekend (Aggenbach, mondelinge mededeling). Soms gaat het om 'noodbloei': als de omstandigheden voor de soort verslechteren, zullen de individuen voordat het te laat is investeren in een sterke groei, bloei en zaadproductie, waarna alle reserves op zijn en de soort verdwijnt. De theorie dat soorten de grootste

bedekking bereiken in het optimum dat door de indicatiewaarde aangeduid wordt, kent dus enkele uitzonderingen. Dit zou een verklaring kunnen zijn voor de korte *BWW*-schaal in de optimale methode voor vochtindicatie.



Figuur 18: De bedekkingswegwaarde-transformatie zoals die door het CML (Groen et al. 1993) gehanteerd wordt in het programma ECOTYP.

De relatief kleine verschillen tussen hoogste en laagste *BWW* bij de vochtindicatie kunnen ook het gevolg zijn van verstoring door andere factoren. Wanneer het vochtgehalte optimaal is voor een soort, en de voedselrijkdom of pH totaal niet, zal de soort geen hoge bedekkingen bereiken. Andersom kan een soort waarvoor het vochtgehalte niet optimaal is, en de voedselrijkdom en zuurgraad wel, wel hogere bedekkingen bereiken (tenzij het vochtgehalte erg extreem is). Volgens dit onderzoek zou de respons van een soort op het vochtgehalte dan veel sterker af moeten hangen van andere factoren dan de respons op pH, want voor de zuurgraadindicatie lijken de maximale verschillen in *BWW* veel groter te zijn. Volgens Kowarik & Seidling (1989) is geen goede indicatie te verkrijgen als een andere dan de te indiceren factor een erg sterke invloed op de vegetatiesamenstelling heeft, bijvoorbeeld op plekken die zeer sterk antropogeen beïnvloed zijn (betreding, maaien). Wanneer bijvoorbeeld betreding een sterk limiterende factor is voor de vegetatie, kunnen alleen soorten die aangepast zijn aan betreding overleven. Op de plek zal de concurrentie tussen soorten beperkt worden door betreding, dus de soorten beperken elkaar niet door concurrentie tot hun ecologisch optimum. Het is dan mogelijk dat soorten onder omstandigheden groeien die verder van hun ecologisch optimum (en dus indicatiewaarde) afwijken dan gewoonlijk. In het algemeen is te stellen dat vermindering van de concurrentie door verstoringen ertoe leidt dat de indicatiewaarden minder goed de situatie ter plekke beschrijven. De verschillen tussen de optimale *BWW*-functies voor vocht- en zuurgraadindicatie kunnen vanuit deze redenering ook het gevolg zijn van het gebruik van verschillende datasets, hoewel dit niet aannemelijk is. De gegevens van Schaffers, gebruikt voor de optimalisatie van de *BWW*-functie bij zuurgraadindicatie, zijn immers verzameld in wegbermen, die veel meer antropogeen beïnvloed zijn dan de natuurgebieden waar de gegevens van Runhaar verzameld zijn. Op grond hiervan zouden de verschillen tussen *BWW*'s voor de zuurgraadindicatie kleiner moeten zijn

dan de verschillen tussen de *BWW*'s voor vochtindicatie (hetgeen niet het geval is). Het blijft echter nuttig om er bij stil te staan dat verschillen in bedekking niet het gevolg hoeven te zijn van de te indiceren factor, en dat de bedekking daarom niet erg belangrijk hoeft te zijn.

De resultaten van het gebruik van Ellenberg F-getallen kunnen verbeterd worden door indicatieweegwaarden (vergelijking (7)) toe te passen, zodat soorten met een meer extreme indicatiewaarde een hoger gewicht krijgen dan soorten die bij gematigd vochtige omstandigheden voorkomen. In deze indicatieweegwaarden wordt gebruik gemaakt van de indifferente soorten, hetgeen de correlatie met de GVG verbetert, aangezien indifferente soorten extreme (in deze studie vooral natte) omstandigheden mijden (zie Figuur 15).

Bij het ecotopensysteem levert een methode waarin de breedte van de ecologische amplitude, afgeleid uit het voorkomen van een soort in één of meer ecologische groepen, gebruikt wordt voor het geven van een gewicht aan afzonderlijke soorten (de soortweegwaardefuncties (12) en (13)) geen significant betere resultaten op (Tabel 29 en Tabel 30). Ook het gebruik van soortspecifieke bedekkingen bij zowel de Ellenberg F-getallen als de indicatiewaarden van het ecotopensysteem zorgde niet voor een significante verbetering (Tabel 23 tot en met Tabel 28). Bij de berekening met soortspecifieke bedekkingen kan het zijn dat er nog niet genoeg gegevens voorhanden waren om van alle soorten een betrouwbare schatting van de soortspecifieke bedekking te krijgen.

Zelinka & Marvan (1961) losten het probleem, dat goede indicatoren overschaduwd worden door nietszeggende soorten, op door aan elke soort een waarde toe te kennen die hoger is naarmate de soort een betere indicator is. Deze waarde werd gebruikt als gewicht voor de indicatiewaarde van de soort. Deze wijze van het toekennen van gewichten aan soorten komt ongeveer overeen met de hier gebruikte soortweegwaarden, die zoals gezegd geen significant effect opleveren. Zelinka & Marvan laten niet zien welk effect hun indicatiegewichten hebben ten opzichte van andere methoden. Ter Braak & Barendregt (1986) voorspellen op theoretische gronden dat een amplitude-gewogen gemiddelde efficiënter is. Met de hier toegepaste amplitude-weging (*SWW*) kon dit effect niet aangetoond worden.

In een studie naar vernatting in de Veenkampen concludeerden Runhaar & Jansen (1999) dat bij Ellenberg vochtgetallen het meenemen van de bedekking bij het berekenen van de (gewogen) gemiddelde indicatiewaarde niet tot een verbetering leidt. Welke *BWW*-methode zij gebruikten is niet beschreven, maar waarschijnlijk is dat de kwantitatieve of decimale methode, zodat de conclusie van Runhaar & Jansen overeen komt met de resultaten van deze studie.

In Bijlage 4 is een aanvullend onderzoek naar de recentelijk door Kiwa Water Research ontwikkelde indicatorsoortenmethode weergegeven. Deze kennisintensieve methode (kennisintensief omdat er met verschillende indicaties voor groepen van vegetatietypen gewerkt wordt, en de indicaties afgeleid zijn uit metingen) presteert niet beter dan de kwalitatieve methode met Ellenberg F-getallen en zelfs slechter dan

de kwalitatieve methode met ecotopensysteem-vochtindicatiewaarden. Blijkbaar is een standplaatsindicatie die gebaseerd is op een select aantal indicatorsoorten niet betrouwbaar of gevoelig genoeg. Ook uit de optimalisaties in deze studie blijkt, dat 'onderwaarderen' van weinig indicatieve soorten niet tot een verbetering leidt, en dat meenemen van indifferente soorten in de analyse wel een positief effect heeft. Het lijkt raadzaam om een indicatiesysteem niet alleen op de duidelijk indicatieve soorten te concentreren.

Om te onderzoeken wat het effect van de onbetrouwbaarheid van de bedekkingsschattingen is, heb ik bij de opnamen van Runhaar de bedekkingsschatting van de (willekeurig geselecteerde) helft van de soorten één schaaleenheid op de schaal van Braun-Blanquet verlaagd. Alleen bij soorten met schatting 'r' (= één exemplaar) kon dit niet, dus wanneer een dergelijke soort geselecteerd was om de bedekking te verlagen, is dit eenvoudigweg niet gebeurd. Op deze manier wordt een andere onderzoeker gesimuleerd, die dezelfde opnamen maakt en bedekkingen iets onderschat, waardoor bedekkingen soms een schaaleenheid lager ingeschat worden. Vervolgens is deze veranderde set opnamen gebruikt in alle permutatietesten die ook met de originele set gegevens uitgevoerd zijn. De conclusies van deze permutatietesten zijn in alle gevallen echter dezelfde als met de originele gegevens, de overschrijdingskansen verschillen hoogstens enkele tienden procenten. Het lijkt vooralsnog dus niet erg veel uit te maken dat de schatting van de bedekking enigszins verschilt tussen onderzoekers. Anders wordt het misschien wanneer gegevens van verschillende onderzoekers bij elkaar genomen worden, of als onderzoekers soorten over het hoofd zien (Lepš & Hadincová 1992). Verder onderzoek naar deze effecten is nodig voor er conclusies getrokken worden.

Misschien speelt in de resultaten mee dat alle opnamen die in deze studie gebruikt zijn gemaakt zijn met de gemodificeerde schaal van Braun-Blanquet. Een andere bedekkingsschaal zorgt er wellicht voor dat onderzoekers op een andere wijze de bedekking inschatten. In ieder geval kan een meer of minder gedifferentieerde schatting van de bedekking effect hebben op de vorm van de optimale *BWW*-functie, omdat de bedekkingen van de soorten dan op andere punten geconcentreerd zijn (nu zijn bijvoorbeeld alle soorten met schatting '4' behandeld als soorten met een bedekking van 62,5%, soorten met bedekking 75% zijn zogenaamd niet aanwezig).

Strikt genomen zijn de resultaten van deze studie alleen geldig voor de opnamen die hier gebruikt zijn. Er is echter weinig reden om aan te nemen dat deze opnamen zo gemaakt zijn dat de conclusies van deze studie hun geldigheid verliezen wanneer ze op andere gegevens toegepast worden. Omdat het onmogelijk is om van de vegetatie een aselechte en willekeurige steekproef te nemen, zal de geldigheid van de hier getrokken conclusies herhaaldelijk getest moeten worden op andere gegevens. Zolang niet bewezen is dat deze conclusies niet voor andere dan de hier gebruikte gegevens van Runhaar en Schaffers gelden, kunnen we aannemen dat ze ook in het algemene geval gelden.

4.3 Belangrijkste conclusies

Door de in dit onderzoek gebruikte permutatietest en bootstrap-schatting kan in veel gevallen toeval van systematisch effect onderscheiden worden, wat de kracht van de uitspraken vergroot. Veel onverwachte resultaten blijken significant te zijn. Hier volgt een korte samenvatting van de belangrijkste resultaten.

- De kwantitatieve methode ($BWW = B$) leidt tot wisselende resultaten, die soms zelfs veel slechter zijn dan de andere methoden. Het gebruik van de kwantitatieve methode is af te raden.
- De resultaten van de kwalitatieve ($BWW = 1$) en decimale methode ($BWW \in \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$) verschillen meestal niet veel van elkaar.
- Voor het gebruik van Ellenberg F-getallen en de vochtindicatiewaarden uit het ecotopensysteem is een optimale bedekkingsweegwaardemethode gevonden, die gekenmerkt wordt door slechts een korte BWW -schaal: soorten met minimale bedekking krijgen $0,4 \times$ gewicht van soorten met maximale bedekking. Formule (18) geeft de optimale methode:

$$BWW_i = 0,4 + 1,5 \left(\frac{B_i - B_{\min}}{B_{\max} - B_{\min}} \right) - 0,9 \left(\frac{B_i - B_{\min}}{B_{\max} - B_{\min}} \right)^3.$$

Het relatief hoge gewicht van soorten die slechts weinig bedekken en de korte BWW -schaal geven aan dat, wanneer het doel van een vegetatieopname is om de vochtindicatie met indicatiewaarden te bepalen, de tijd het beste gestoken kan worden in het zoeken naar eventueel nog ontbrekende soorten dan in nauwkeurige schattingen van de bedekkingen.

- Bij de zuurgraadindicatie is er nauwelijks significant verschil tussen alle onderzochte methoden. Het kan zijn dat de zuurgraadindicatie daadwerkelijk ongevoelig is voor de transformaties van bedekkingen in gewichten, of dat de gebruikte steekproef te klein is.
- Het rekenen met gewichten die afhangen van de soortspecifieke bedekking (gemiddelde of maximale bedekking, formule (5)) van een soort heeft geen effect.
- Soorten kunnen in het ecotopensysteem een lager gewicht krijgen naarmate hun ecologische amplitude breder is en hun 'indicatiekracht' kleiner zou moeten zijn. Het toepassen van dergelijke soortweegwaarden (berekend volgens formules (12) en (13)) leidt echter niet tot een verbetering.

- De berekening van gemiddelde vochtindicatiewaarden met Ellenberg F-getallen kan significant verbeterd worden door indicatieweegwaarden toe te kennen aan de indicatiewaarden, afhankelijk van de indifferente soorten in de opname (formule (7)):

$$IWW_i = 1 + c_1 \times \left(1 - \frac{c_2 \left(\sum BWW_{\text{indif}} \right)}{1 + c_2 \left(\sum BWW_{\text{indif}} \right)} \right) \times \left(\frac{I_c - I}{I_c - I_{\text{min}}} \right)^2.$$

Voor optimaal resultaat moeten de waarden van de parameters c_1 en c_2 hoog zijn, tussen 20 en 70. Op deze manier krijgen soorten van extreme milieus een veel hogere indicatiewaarden dan soorten van meer gematigde milieus, wanneer er geen indifferente soorten aanwezig zijn. Zijn er wel indifferente soorten, dan nemen de gewichten van de extreme indicatiewaarden snel af. Indifferente soorten komen, door het ontbreken van specialistische aanpassingen aan extreme omstandigheden, voor in meer gematigde milieus. We kunnen indifferente soorten dus als indicatoren voor gematigde milieus beschouwen. Het lijkt vanuit dit resultaat van belang dat een indicatiesysteem informatie over zoveel mogelijk soorten in de uitspraak betreft.

Literatuur

- Aggenbach, C.J.S., Jalink, M.H. & Jansen, A.J.M. (1998) Indicatorsoorten voor verdroging, verzuring en eutrofiëring van plantengemeenschappen in vennen. Deel 5 uit de serie 'Indicatorsoorten'. Staatsbosbeheer, Driebergen
- Alkemade, J.R.M., Wiertz, J. & Latour, J.B. (1996) Kalibratie van Ellenbergs milieu-indicatiegetallen aan werkelijk gemeten bodemfactoren. RIVM, Bilthoven
- Auerbach, M. & Shmida, A. (1993) Vegetation change along an altitudinal gradient on Mt Hermon, Israel – no evidence for discrete communities – *Journal of Ecology* **81**, 25-33
- Böcker, R., Kowarik, I. & Bornkamm, R. (1983) Untersuchungen zur Anwendung der Zeigerwerte nach Ellenberg – *Verhandlungen der Gesellschaft für Ökologie (Festschrift Ellenberg)* **11**, 35-56
- Clausman, P.H.M.A. & Van Wijngaarden, W. (1984) Verspreiding en ecologie van wilde planten in Zuid-Holland; 1a waarderingsparameters. Provinciale Planologische Dienst Zuid-Holland
- Diaconis, P. & Efron, B. (1983) Computer-intensive methods in statistics – *Scientific American* **248**, 116-130
- Durwen, K.-J. (1982) Zur Nutzung von Zeigerwerten und artspezifischen Merkmalen der Gefäßpflanzen Mitteleuropas für Zwecke der Landschaftsökologie und -planung mit Hilfe der EDV – Voraussetzungen, Instrumentarien, Methoden und Möglichkeiten – . Arbeitsberichte des Lehrstuhls Landschaftsökologie Münster, Heft 5, Münster
- Durwen, K.-J. (1983) Bioindikation im Dienste des Umweltschutzes – *Beiträge zur Landespflege in Rheinland-Pfalz* **9**, 133-160
- Edgington, E.S. (1995) Randomization tests. 3rd edition, revised and expanded. Statistics: Textbooks and Monographs. Marcel Dekker, New York
- Efron, B. & Tibshirani, R.J. (1993) An Introduction to the Bootstrap. Monographs on Statistics and Applied Probability 57, Chapman & Hall, New York
- Ellenberg, H. (1974) Zeigerwerte der Gefäßpflanzen Mitteleuropas. Scripta Geobotanica IX, Verlag Erich Goltze, Göttingen
- Ellenberg, H. (1992) Zeigerwerte der Gefäßpflanzen (ohne *Rubus*), in: Ellenberg, H., Weber, H.E., Düll, R., Wirth, V., Werner, W. & Paulßen, D., Zeigerwerte von Pflanzen in Mitteleuropa, 2. und verbesserte Auflage. Scripta Geobotanica XVIII, Verlag Erich Goltze, Göttingen
- Groen, K., Stevers, R., Van Gool, C. & Broekmeyer, M. (1993) Uitwerking ecotopensysteem fase III. Herziene landelijke typologie en vertaalsleutels voor Overijssel, Gelderland, Noord-Brabant en Limburg. CML mededelingen 49, Centrum voor Milieukunde, Leiden
- Hertog, A.J. & Rijken, M. (1992) Geautomatiseerde bepaling van natuurbehoudswaarde in vegetatieopnamen. Provincie Gelderland, Dienst Ruimte, Wonen en Groen

- Hill, M.O., Roy, D.B., Mountford, J.O. & Bunce, R.G.H. (2000) Extending Ellenberg's indicator values to a new area: an algorithmic approach – *Journal of Applied Ecology* **37**, 3-15
- Hjorth, J.S.U. (1994) Computer intensive statistical methods; validation, model selection and bootstrap. Chapman & Hall, London
- Jansen, P.C., Runhaar, J., Witte, J.P.M. & Van Dam, J.C. (2000) Vochtindicatie van graslandvegetaties in relatie tot de vochttoestand van de bodem. Alterra-rapport 057, Alterra & Wageningen Universiteit, Wageningen
- KNMI (1990) Het weer, jaar 1990 – *Jaaroverzicht van het weer in Nederland* **87** (13)
- KNMI (1991) Jow, jaar 1991 – *Jaaroverzicht van het weer in Nederland* **88** (13)
- KNMI (1992) Jow, jaar 1992 – *Jaaroverzicht van het weer in Nederland* **89** (13)
- KNMI (1993) Jow, jaar 1993 – *Jaaroverzicht van het weer in Nederland* **90** (13)
- KNMI (1994) Jaar 1994 – *Jaaroverzicht van het weer in Nederland* **91** (13)
- Knotters, M. & Bierkens, M.F.P. (1999) Hoe lang moet je de grondwaterstand meten om iets over de dynamiek te weten? – *Stromingen* **5** (4), 5-12
- Kowarik, I. & Seidling, W. (1989) Zeigerwertberechnungen nach Ellenberg – Zu Problemen und Einschränkungen einer sinnvollen Methode – *Landschaft + Stadt* **21** (4), 132-143
- Kunzmann, G. (1989) Der ökologische Feuchtegrad als Kriterium zur Beurteilung von Grünlandstandorten, ein vergleich bodenkundlicher und vegetationskundlicher Standortmerkmale. Dissertationes Botanicae 134
- Lepš, J. & Hadincová, V. (1992) How reliable are our vegetation analyses? – *Journal of Vegetation Science* **3**, 119-124
- Lienert, G.A. (1986) Verteilungsfreie Methoden in der Biostatistik, Band 1. 3. Auflage, Verlag Anton Hain, Meisenheim am Glan
- Lienert, G.A. (1975) Verteilungsfreie Methoden in der Biostatistik, Tafelband. Verlag Anton Hain, Meisenheim am Glan
- Möller, H. (1987) Wege zur Ansprache der aktuellen Bodenazidität auf der Basis der Reaktionszahlen von Ellenberg ohne arithmetisches Mitteln dieser Werte – *Tuexenia* **7**, 499-505
- Mueller-Dombois, D. & Ellenberg, H. (1974) Aims and methods of vegetation ecology. John Wiley & Sons, New York
- Noest, V., Van der Maarel, E., Van der Meulen, F. & Van der Laan, D. (1989) Optimum-transformation of plant species cover-abundance values – *Vegetatio* **83**, 167-178
- Runhaar, J. (1989a) Toetsing van het ecotopensysteem I: hoofd rapport. CML mededelingen 48a, Centrum voor Milieukunde, Leiden
- Runhaar, J. (1989b) Toetsing van het ecotopensysteem II: rapportage van het veldwerk. CML mededelingen 48b, Centrum voor Milieukunde, Leiden
- Runhaar, J. & Jansen, P.C. (1999) Standaard meetprotocol verdroging; Vegetatiemonitoring. Nationaal Onderzoeksprogramma Verdroging, rapport 15-3. RIZA, Lelystad
- Runhaar, J., Witte, J.P.M. & Verburg, P.H. (1997) Ground-water level, moisture supply, and vegetation in the Netherlands – *Wetlands* **17** (4), 528-538
- Sachs, L. (1984) Angewandte Statistik. Anwendung statistischer Methoden, 6. Auflage. Springer-Verlag, Berlin

- Schaffers, A.P. (2000) Ecology of roadside plant communities. PhD thesis, Wageningen University, Wageningen.
- Schaffers, A.P. & Sýkora, K.V. (2000) Reliability of Ellenberg indicator values for moisture, nitrogen and soil reaction: a comparison with field measurements – *Journal of Vegetation Science* **11**, 225-244
- Schaminée, J.H.J., Stortelder, A.H.F. & Westhoff, V. (1995) De vegetatie van Nederland deel 1: grondslagen, methoden, toepassingen. Opulus Press, Uppsala, Leiden
- Shimwell, D.W. (1971) Description & classification of vegetation. Sidgwick & Jackson, London
- Stevens, R.A.M., Runhaar, J. & Groen, C.L.G. (1987) Het CML-ecotopensysteem. Uitwerking voor Noord-, West- en Zuidwest-Nederland. CML mededelingen 34, Centrum voor Milieukunde, Leiden
- Ter Braak, C.J.F. & Barendregt, L.G. (1986) Weighted averaging of species indicator values: its efficiency in environmental calibration – *Mathematical Biosciences* **78**, 57-72
- Van der Wouw, M. (2000) De GHG, zo veranderlijk als het weer – *Stromingen* **6** (3), 23-28
- Witte, J.P.M. & Runhaar, H. (2000) Planten als indicatoren voor water – *Stromingen* **6** (1), 5-21
- Witte, J.P.M. (2002) The descriptive capacity of ecological plant species groups – *Plant Ecology* **162**, 199-213
- Witte, J.P.M. & Von Asmuth, J.R. (in prep.) Do we really need phytosociological classes to calibrate Ellenberg's indicator values? Ingediend bij: *Journal of Vegetation Science*
- Zelinka, M. & Marvan, P. (1961) Zur Präzisierung der biologischen Klassifikation der Reinheit fließender Gewässer – *Archiv für Hydrobiologie* **57** (3), 389-407

Bijlage 1 Overzicht van veel gebruikte formules

Formule 1 voor het berekenen van **gewogen gemiddelde indicatiewaarden**:

$$I_{\text{gem}} = \frac{\sum_{i=1}^N I_i \times BWW_i}{\sum_{i=1}^N BWW_i}$$

waarin

I_{gem} = gewogen gemiddelde indicatiewaarde;

i = de i^{de} soort in de opname;

N = aantal soorten in de opname;

I_i = indicatiewaarde van soort i ;

BWW_i = bedekkingsweegwaarde van soort i gebaseerd op de bedekking van soort i .

Formule 4, bedekkingsweegwaardefunctie:

$$BWW_i = a_0 + a_1 \left(\frac{B_i - B_{\min}}{B_{\max} - B_{\min}} \right) + (1 - a_0 - a_1) \left(\frac{B_i - B_{\min}}{B_{\max} - B_{\min}} \right)^{a_2}$$

met

BWW_i = bedekkingsweegwaarde van soort i ;

B_i = bedekking van soort i ;

B_{\max}, B_{\min} = maximale respectievelijk minimale bedekking;

a_0, a_1, a_2 = parameters.

Formule 5, soortspecifieke bedekkingsweegwaardefunctie:

$$BWW_i = b_0 + (1 - b_0) \left(\frac{B_i - B_{\min}}{SSB_i - B_{\min}} \right)^{b_1}$$

waarin

BWW_i = bedekkingsweegwaarde van soort i ;

B_i = bedekking van soort i ;

B_{\max}, B_{\min} = maximale respectievelijk minimale bedekking;

SSB_i = specifieke bedekking van soort i ;

b_0, b_1 = parameters.

Formule 7, indicatieweegwaardefunctie:

$$IWW_I = 1 + c_1 \times \left(1 - \frac{c_2 \left(\sum BWW_{\text{indif}} \right)}{1 + c_2 \left(\sum BWW_{\text{indif}} \right)} \right) \times \left(\frac{I_c - I}{I_c - I_{\min}} \right)^2$$

met:

- IWW_I = indicatieweegwaarde van indicatiewaarde I ;
- I = indicatiewaarde, $1 \leq I \leq 9$ of $1 \leq I \leq 12$;
- I_{\min} = kleinste mogelijke indicatiewaarde (= 1);
- I_c = middelste indicatiewaarde, $I_c = 5$ of $I_c = 6,5$;
- c_1, c_2 = parameters.

Formule 8 voor het berekenen van een **gemiddelde indicatiewaarde** met bedekkingsweegwaarden en indicatieweegwaarden:

$$I_{\text{gem}} = \frac{\sum_{i=1}^N I_i \times BWW_i \times IWW_{I_i}}{\sum_{i=1}^N BWW_i \times IWW_{I_i}}$$

waarin

- I_{gem} = gewogen gemiddelde indicatiewaarde;
- i = de i^{de} soort in de opname;
- N = aantal soorten in de opname;
- I_i = indicatiewaarde van soort i ;
- BWW_i = bedekkingsweegwaarde van soort i gebaseerd op de bedekking van soort i ;
- IWW_{I_i} = indicatieweegwaarde van de indicatiewaarde I_i van soort i .

Formule 9 voor het bepalen van de **indicatiewaarde van het ecotopensysteem** uit de verdeling van de soorten over de klassen:

$$I_i = \sum_{j=1}^M F_{i,j} \times j$$

waarin

- I_i = indicatiewaarde van soort i ;
- j = klasse van het ecotopensysteem;
- $F_{i,j}$ = de weegfactor van klasse j van soort i ;
- M = het aantal indicatieklassen voor de betreffende milieufactor.

Formule 11 voor het berekenen van **gemiddelde indicatiewaarden** met behulp van het ecotopensysteem, met bedekkingsweegwaarden:

$$I_{\text{gem}} = \frac{\sum_{i=1}^N \left(\sum_{j=1}^M F_{i,j} \times j \right) \times BWW_i}{\sum_{i=1}^N BWW_i}$$

waarin

- I_{gem} = gewogen gemiddelde indicatiewaarde;
- i = de i^{de} soort in de opname;
- N = aantal soorten in de opname;
- I_i = indicatiewaarde van soort i ;
- BWW_i = bedekkingsweegwaarde van soort i gebaseerd op de bedekking van soort i ;
- j = klasse van het ecotopensysteem;
- $F_{i,j}$ = de weegfactor van klasse j van soort i ;
- M = het aantal indicatieklassen voor de betreffende milieufactor.

Formule 12:

$$SWW_i = \sum_{j=1}^M F_{i,j}^d \quad \text{voor } d \geq 1$$

en **formule 13, soortweegwaardefuncties:**

$$SWW_i = \frac{1}{\sum_{j=1}^M F_{i,j}^d} \quad \text{voor } 0 \leq d \leq 1$$

waarin

- SWW_i = soortweegwaarde van soort i ;
- $F_{i,j}$ = de weegfactor van klasse j van soort i ;
- M = het aantal indicatieklassen voor de betreffende milieufactor;
- d = parameter.

Formule 14 voor het berekenen van **gemiddelde indicatiewaarden** met behulp van het ecotopensysteem, met bedekkingsweegwaarden en soortweegwaarden:

$$I_{\text{gem}} = \frac{\sum_{i=1}^N \left(\sum_{j=1}^M F_{i,j} \times j \right) \times BWW_i \times SWW_i}{\sum_{i=1}^N BWW_i \times SWW_i}$$

waarin

- I_{gem} = gewogen gemiddelde indicatiewaarde;
- i = de i^{de} soort in de opname;
- N = aantal soorten in de opname;
- I_i = indicatiewaarde van soort i ;
- SWW_i = soortweegwaarde van soort i ;
- BWW_i = bedekkingsweegwaarde van soort i ;
- j = klasse van het ecotopensysteem;
- F_{ij} = de weegfactor van klasse j van soort i ;
- M = het aantal indicatieklassen voor de betreffende milieufactor.

Formule 15 voor het berekenen van de **mediaan** van waarnemingen in klassen:

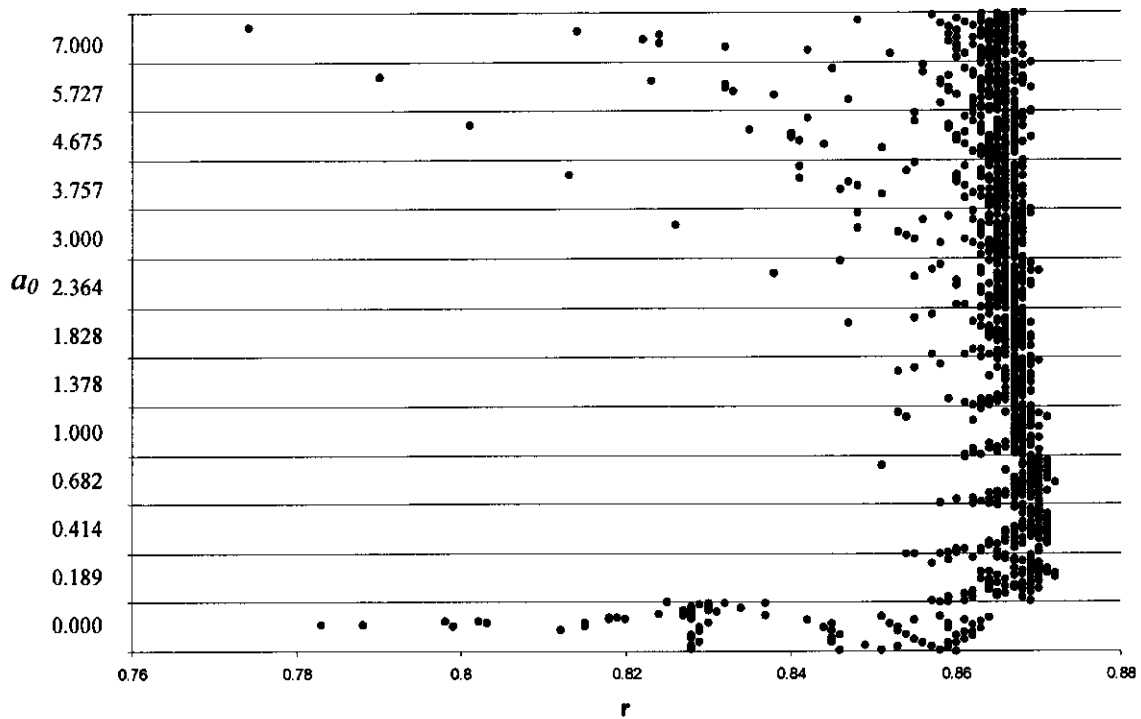
$$I_m = O + b \times \left(\frac{(N/2) - N_o}{N_i} \right)$$

met

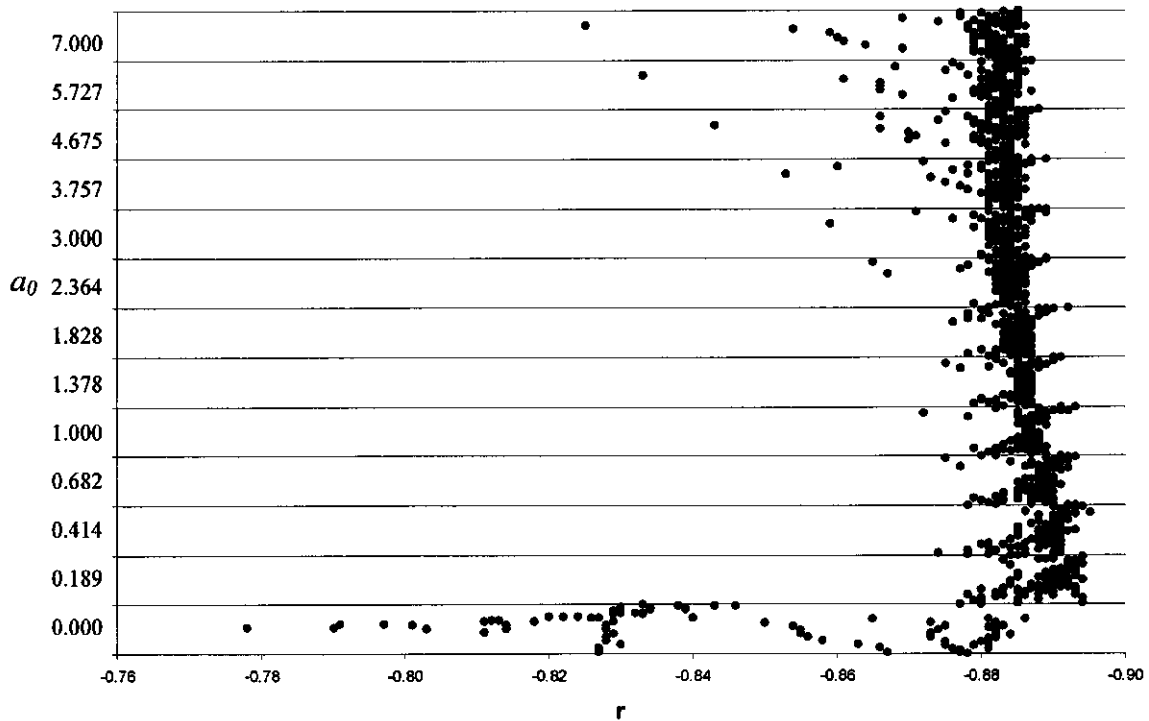
- I_m = mediane indicatiewaarde;
- O = de ondergrens van de klasse waarin de middelste waarneming valt;
- b = de breedte van de klasse waarin de middelste waarneming valt;
- N = het aantal soorten per opname;
- N_o = het aantal soorten onder de klasse waarin de middelste waarneming valt;
- N_i = het aantal soorten in dezelfde klasse als de middelste waarneming.

Bijlage 2 Gevoeligheid van de *BWW*-formule voor de parameters.

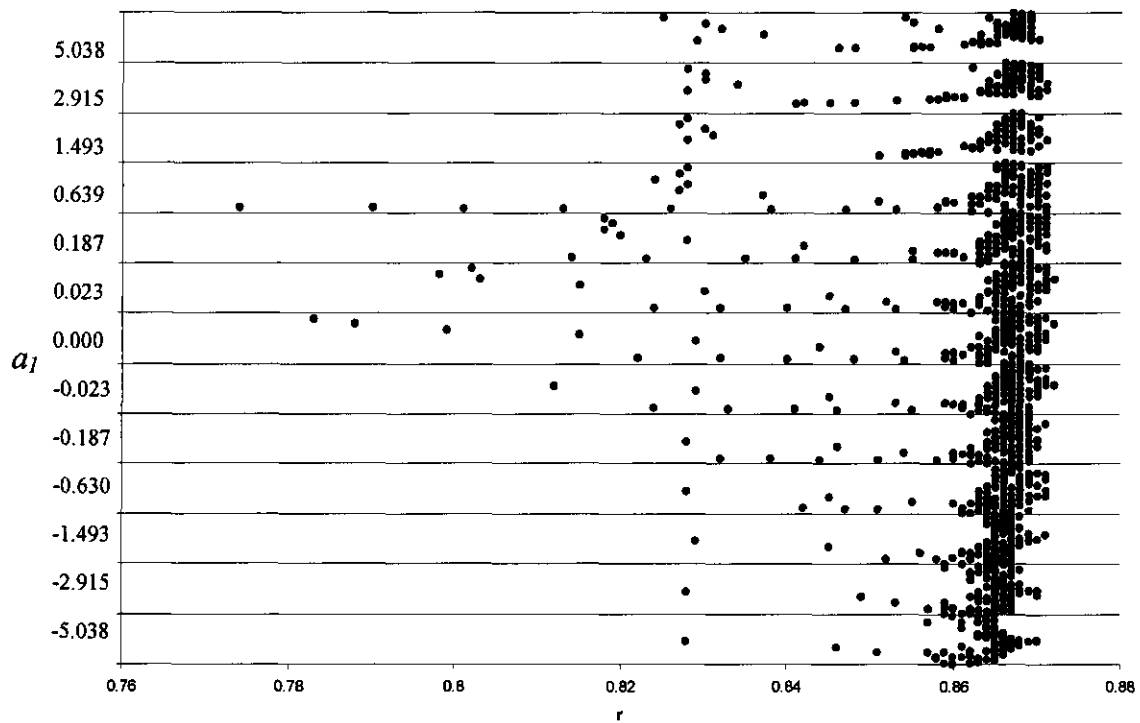
In elke grafiek op de volgende pagina's zijn de correlatiecoëfficiënten van gemiddelde indicatiewaarde met de bijbehorende milieufactor weergegeven voor alle onderzochte *BWW*-functies (dus voor alle waarden van de parameters a_0 , a_1 en a_2). Om het effect van de afzonderlijke parameters te laten zien, zijn de *BWW*-functies gesorteerd, allereerst op de waarden van de parameter die langs de y-as aangegeven is. Binnen elke waarde van deze 'sorteer-parameter' is gesorteerd naar de waarden van de volgende parameter, en binnen elke waarde van deze parameter is weer gesorteerd naar de waarden van de laatste parameter. De volgorde van sorteren (de prioriteiten van de parameters) en de gebruikte gegevens zijn in het onderschrift van elke grafiek gegeven. De Spearman rangcorrelatiecoëfficiënt wordt onder elke grafiek aangeduid met ' r '.



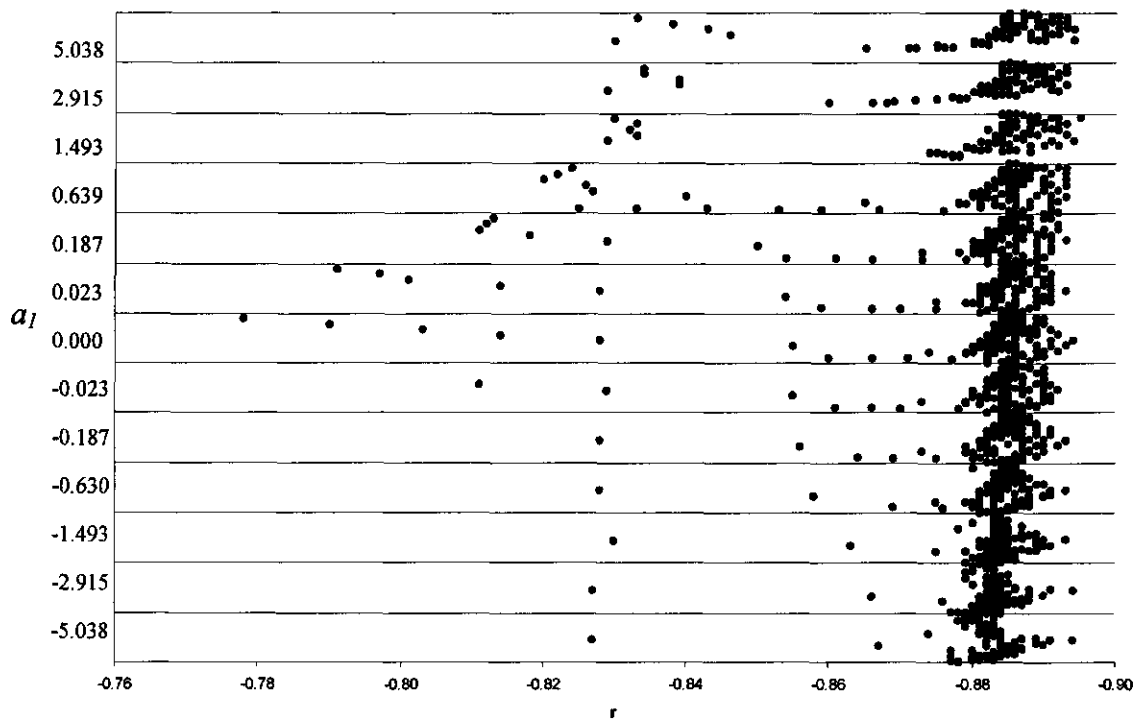
Figuur 19: Correlaties van alle BWW-functies gerangschikt naar a_0 , dan naar a_1 en a_2 . Gegevens: Ellenberg F-getallen en Runhaar GVG.



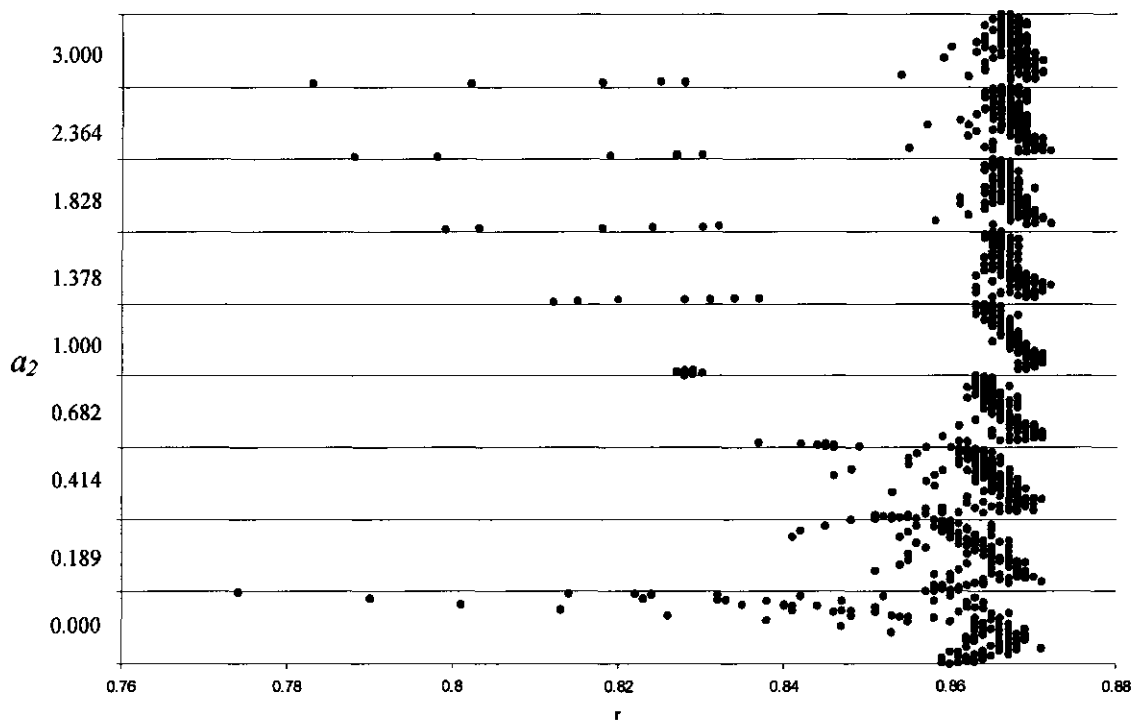
Figuur 20: Correlaties van alle BWW-functies gerangschikt naar a_0 , dan naar a_1 en a_2 . Gegevens: Vochtindicatie volgens het ecotopensysteem en Runhaar GVG.



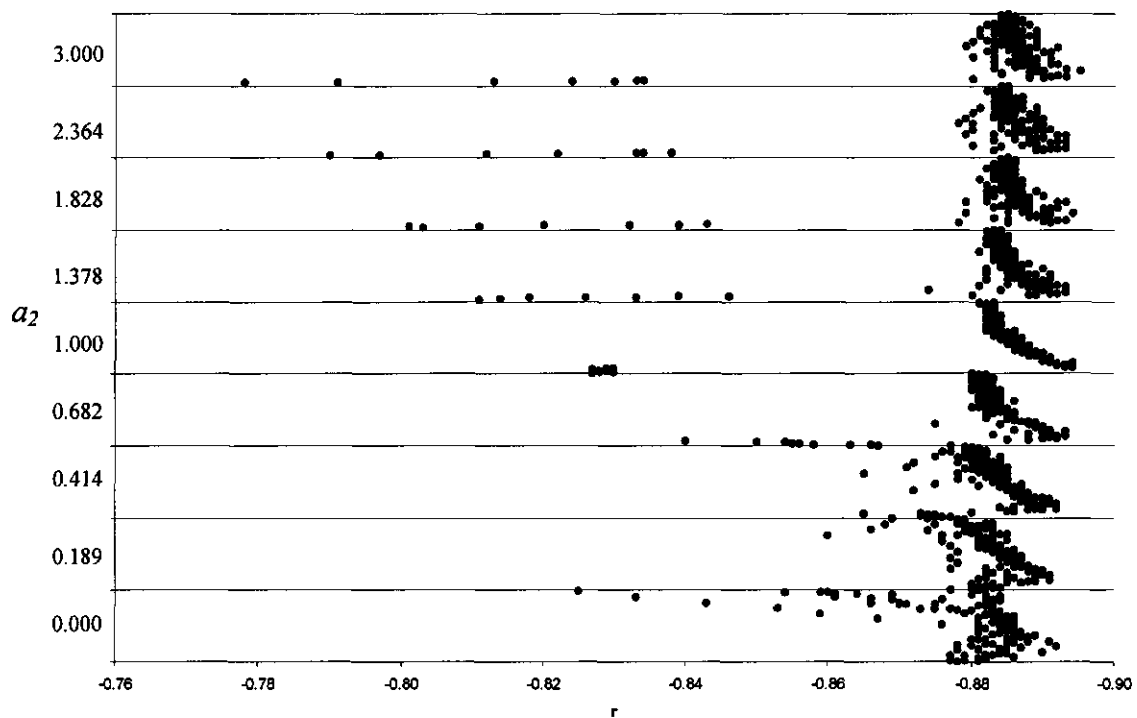
Figuur 21: *Correlaties van alle BWW-functies gerangschikt naar a_1 , dan naar a_2 en a_0 . Gegevens: Ellenberg F-getallen en Runhaar GVG.*



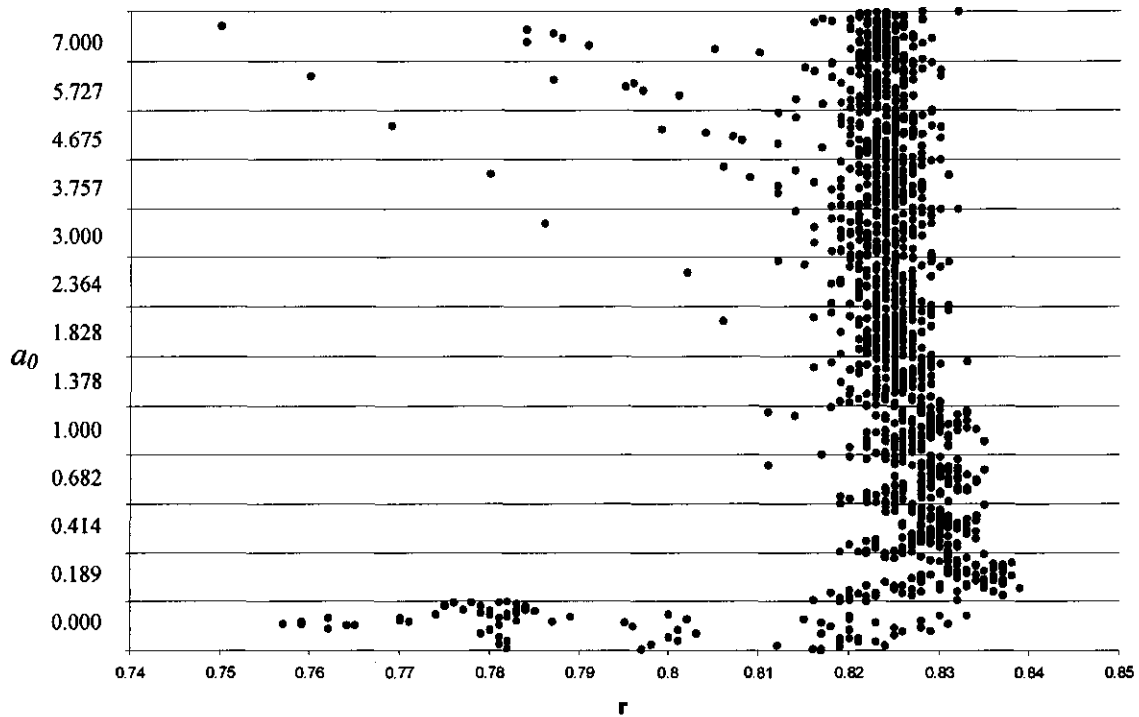
Figuur 22: *Correlaties van alle BWW-functies gerangschikt naar a_1 , dan naar a_2 en a_0 . Gegevens: Vochtindicatie volgens ecotopensysteem en Runhaar GVG.*



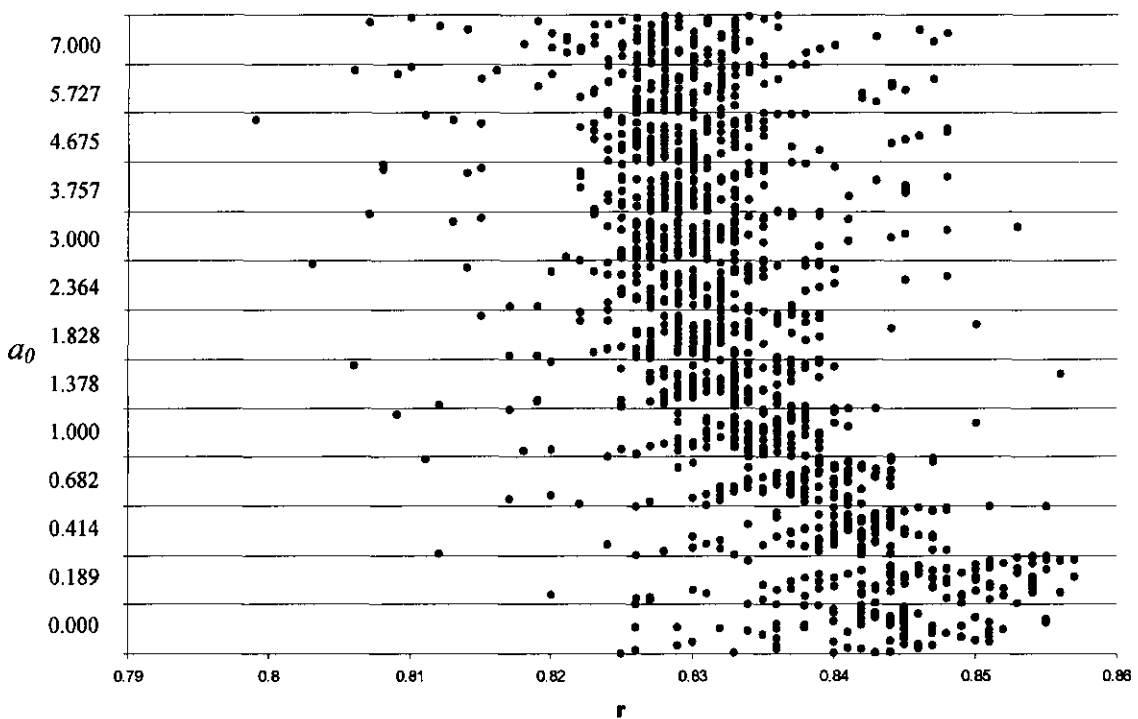
Figuur 23: Correlaties van alle BWW-functies, gerangschikt naar a_2 , dan naar a_0 en a_1 . Gegevens Ellenberg F-getallen en Runhaar GVG.



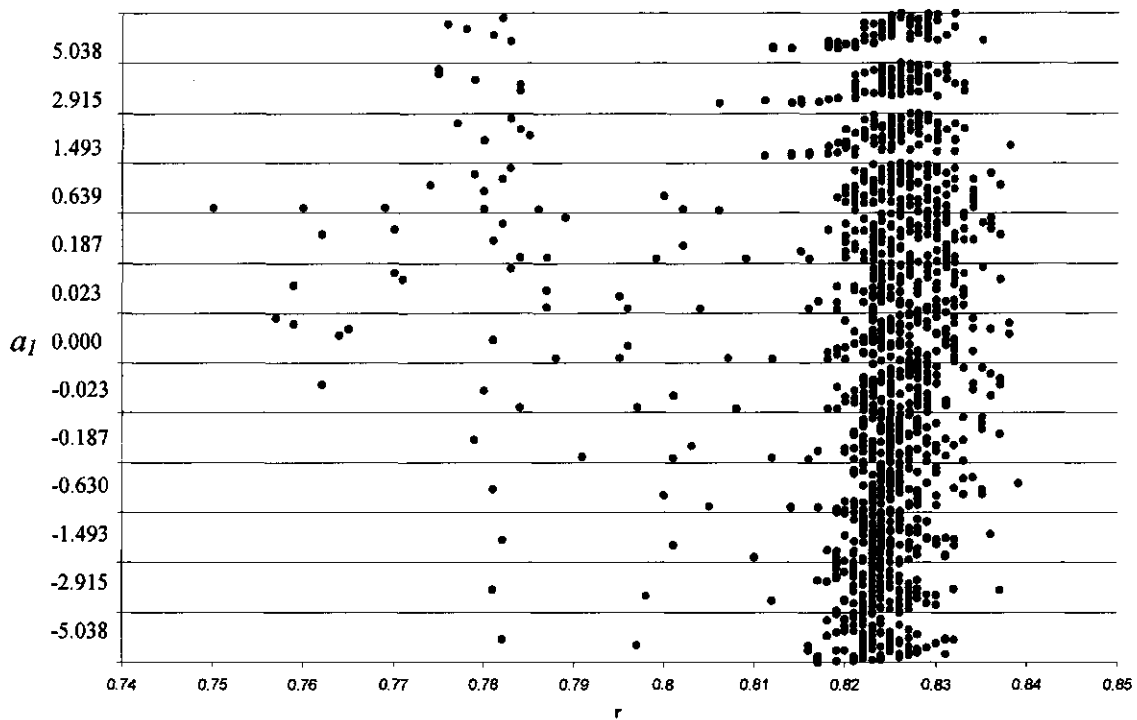
Figuur 24: Correlaties van alle BWW-functies, gerangschikt naar a_2 , dan naar a_0 en a_1 . Gegevens: vochtindicatie volgens ecotopensysteem en Runhaar GVG.



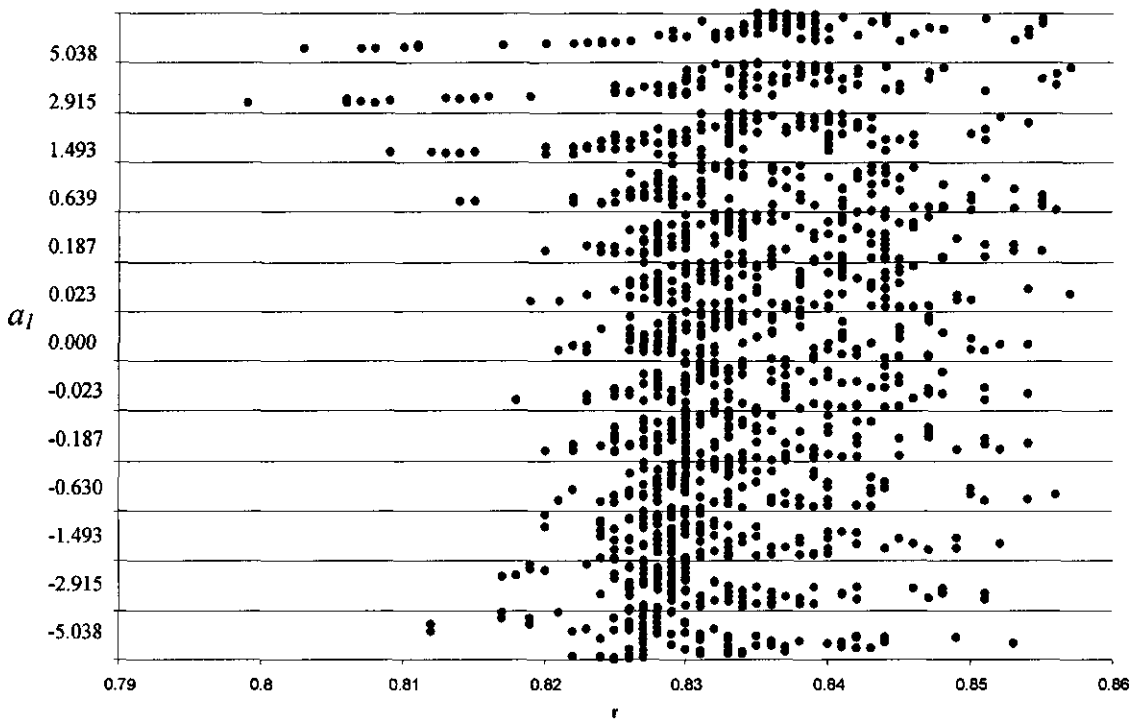
Figuur 25: Correlaties van alle BWV-functies gerangschikt naar a_0 , dan naar a_1 en a_2 . Gegevens: Ellenberg R-getallen en Schaffers pH.



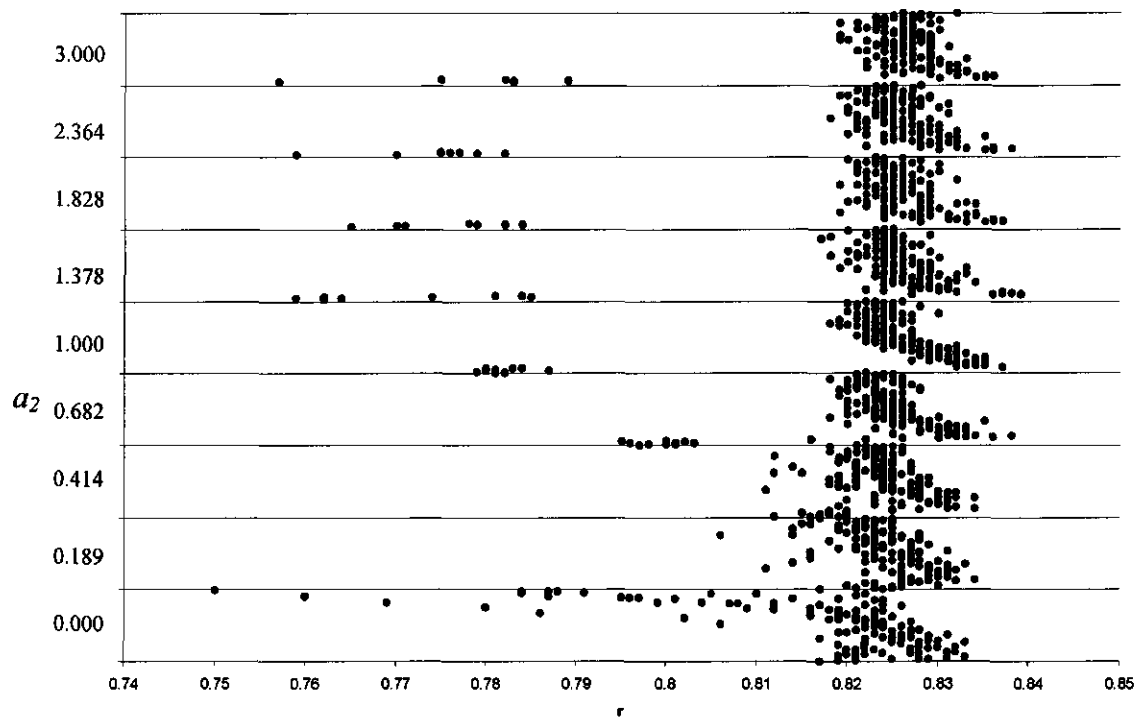
Figuur 26: Correlaties van alle BWV-functies gerangschikt naar a_0 , dan naar a_1 en a_2 . Gegevens: zuurgraadindicatie ecotopensysteem en Schaffers pH.



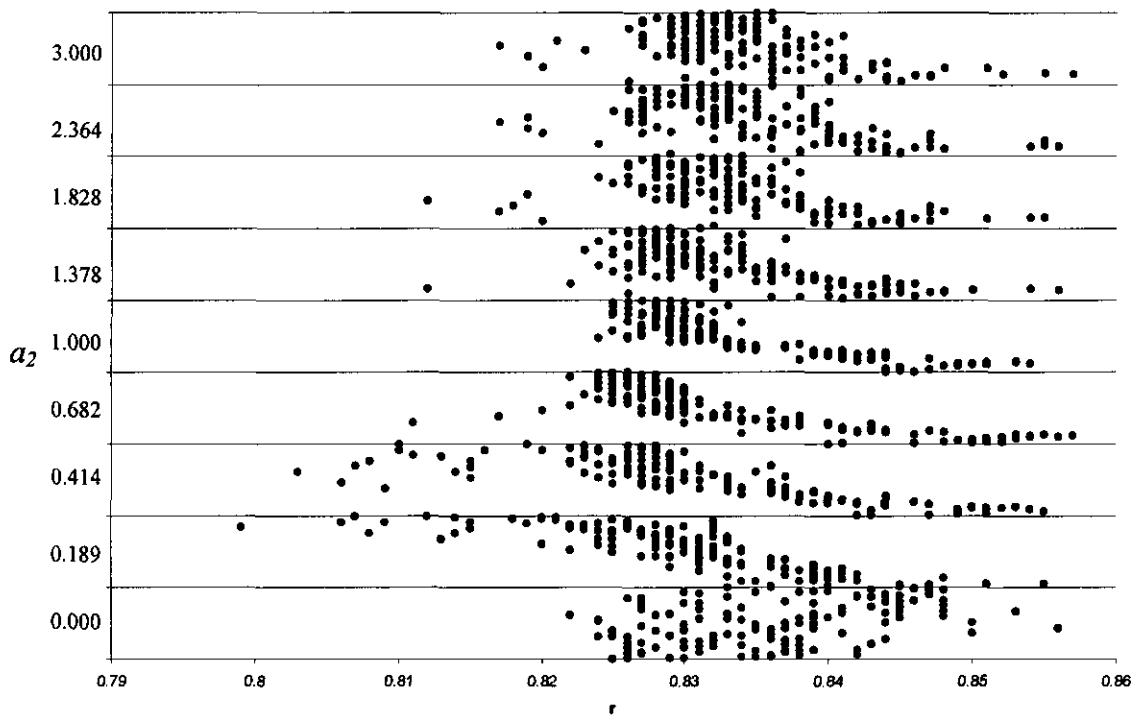
Figuur 27: Correlaties van alle BWW-functies gerangschikt naar a_1 , dan naar a_2 en a_0 . Gegevens: Ellenberg R-getallen en Schaffers pH.



Figuur 28: Correlaties van alle BWW-functies gerangschikt naar a_1 , dan naar a_2 en a_0 . Gegevens: zuurgraadindicatie ecotopensysteem en Schaffers pH.



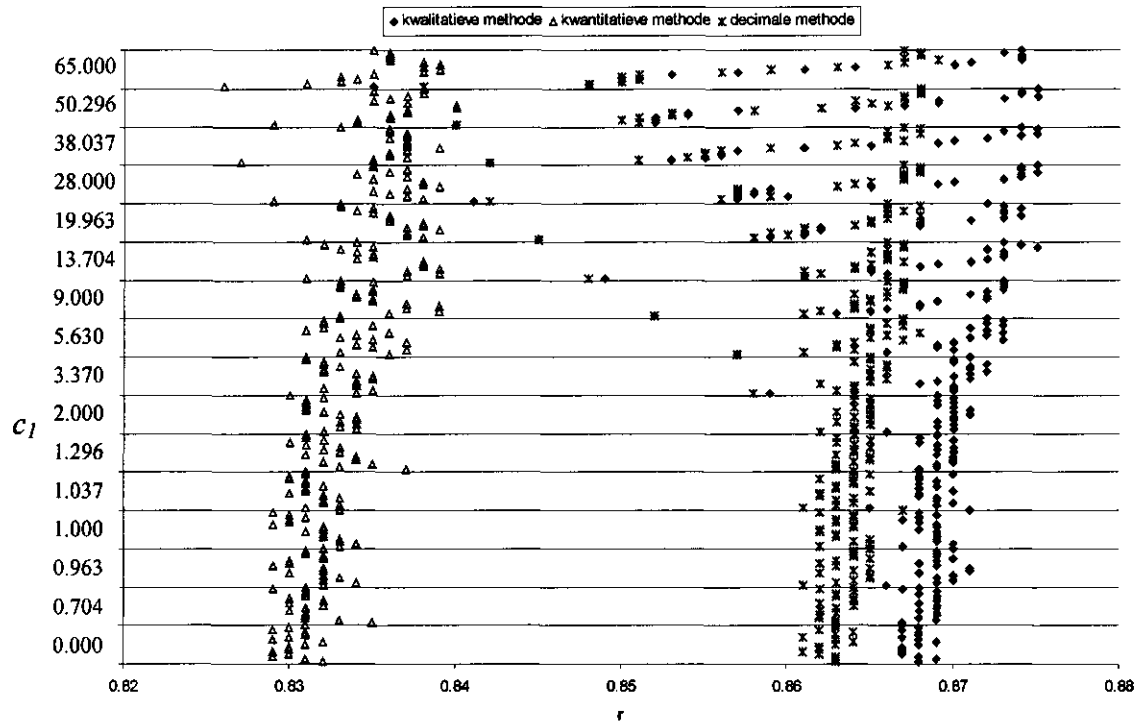
Figuur 29: Correlaties van alle BWW-functies gerangschikt naar a_2 , dan naar a_0 en a_1 . Gegevens: Ellenberg R-getallen en Schaffers pH.



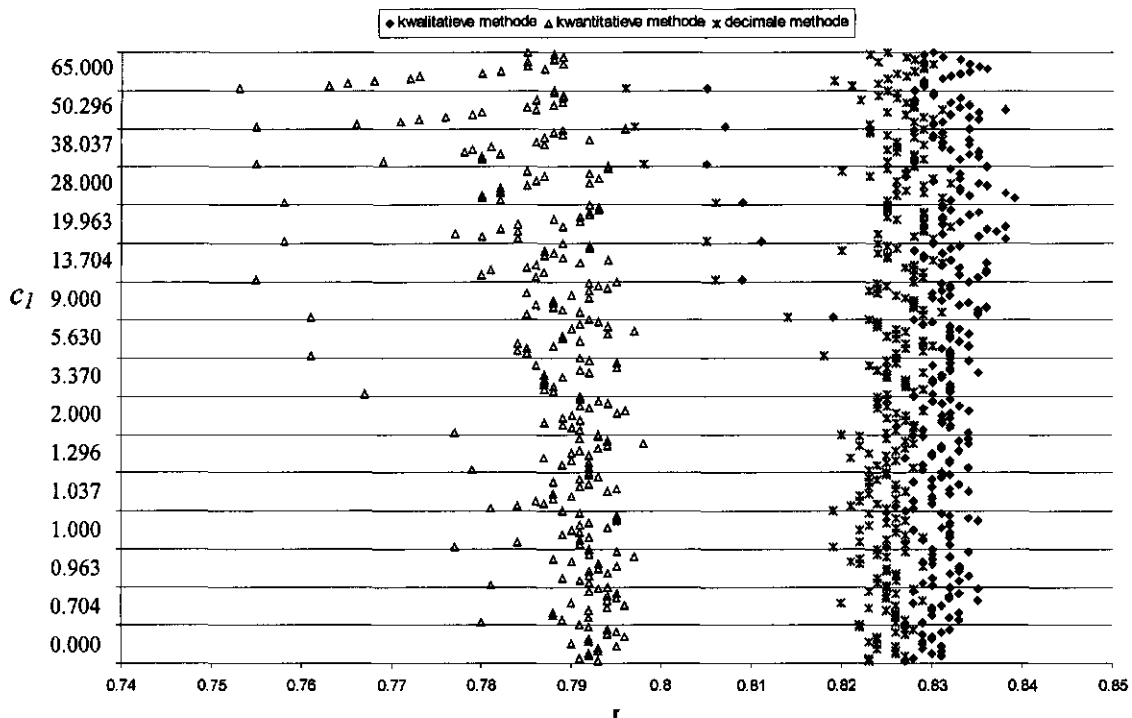
Figuur 30: Correlaties van alle BWW-functies gerangschikt naar a_2 , dan naar a_0 en a_1 . Gegevens: zuurgraadindicatie ecotopensysteem en Schaffers pH

Bijlage 3 Gevoeligheid van de *IWW*-formule voor de parameters.

Op de volgende pagina staan grafieken met de correlaties tussen gemiddelde indicatiewaarde en milieufactoren voor alle combinaties van parameters in de *IWW*-formule. De gegevens zijn gesorteerd naar waarden van de parameter c_0 . Alle punten tussen twee horizontale lijnen zijn berekend met dezelfde waarde van c_1 (die weergegeven is langs de y-as), terwijl de waarde van c_2 varieert (van laag naar hoog). Er is steeds met drie verschillende bedekkingsweegmethoden gerekend. Onder elke grafiek is aangegeven welke indicatiewaarden en milieugegevens gebruikt zijn.



Figuur 31: Correlatie van gemiddelde Ellenberg F-getallen volgens alle IWW-functies en drie BWW-methoden met GVG.



Figuur 32: Correlatie van gemiddelde Ellenberg R-getallen volgens alle IWW-functies en drie BWW-methoden met pH.

Bijlage 4 Standplaatsindicatie met het Kiwa-indicatorsoortensysteem

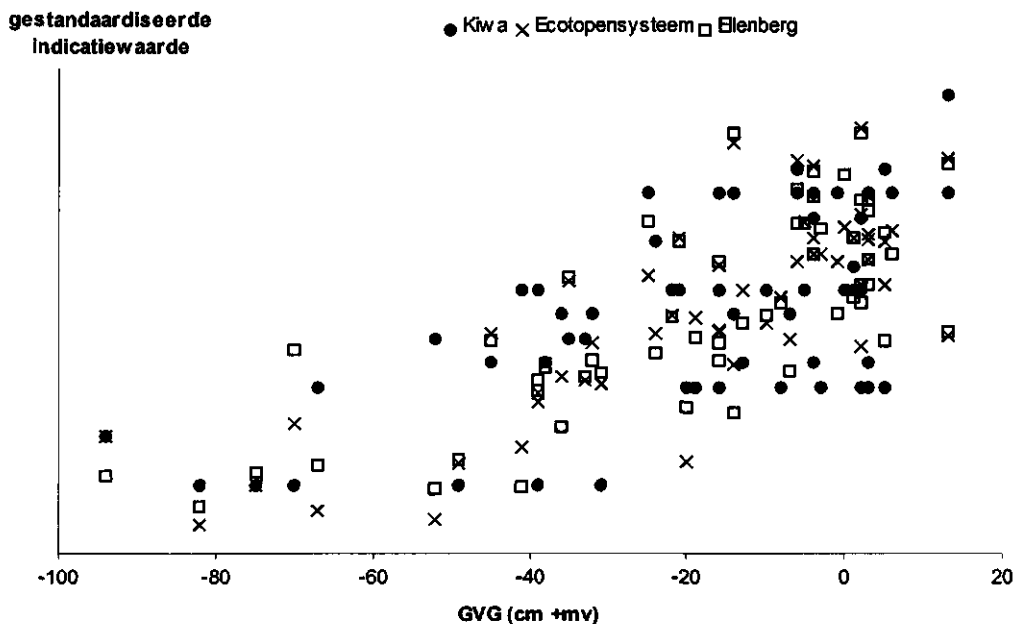
Behalve de indicatiesystemen van Ellenberg en het ecotopensysteem is nog een derde, recentelijk ontwikkeld systeem onderzocht, dat hier aangeduid wordt als 'Kiwa-indicatorsoorten'.

Dit systeem is door het onderzoeksinstituut Kiwa Water Research ontwikkeld. In dit systeem hebben een aantal karakteristieke soorten een indicatiewaarde gekregen, die afhankelijk is van het landschapstype en vegetatietype. De werkwijze om een indicatie te verkrijgen van de standplaats is, kort omschreven, als volgt. De gebruiker moet eerst het landschapstype bepalen, dan het vegetatietype (plantensociologische eenheid). Vervolgens wordt met een tabel behorende bij het landschaps- en vegetatietype per aanwezige soort in de vegetatieopname opgezocht of het een indicatorsoort is, en welke omstandigheden deze soort indiceert. De indicatie van de hele opname is dan ruwweg de indicatieklasse die het vaakst door de afzonderlijke soorten geïndiceerd is (de modaal), en kan ook meerdere indicatieklassen omvatten (voor details, zie onder andere Aggenbach et al. 1998).

Deze procedure was op het moment van onderzoek slechts voor enkele landschapstypen grotendeels geautomatiseerd (op het bepalen van het landschapstype en enkele keuzes voor de bepaling van het vegetatietype na). Om de vochtindicatie volgens de Kiwa-methode te vergelijken met Ellenberg en het ecotopensysteem, is van de opnamen van Runhaar uit de landschapstypen beekdalen, vennen en duinvalleien in de kalkarme duinen (56 opnamen) met de kiwa-methode de vochtindicatie bepaald en gecorreleerd met de grondwaterstand. Dit leverde een r_s van -0,537 op. Deze correlatie verschilt niet significant van de correlatie wanneer Ellenberg F-getallen gebruikt worden ($r_s = 0,868$, $p = 13,0\%$ (tweezijdig)), maar er is wel een significant verschil vergeleken met het ecotopensysteem ($r_s = -0,887$, $p = 4,8\%$ (tweezijdig)), zie Figuur 33. Omdat het kiwa-systeem eigenlijk de grondwaterstand over het hele jaar aanduidt, zijn alle testen ook met de gemiddelde grondwaterstand over het hele jaar uitgevoerd. De conclusies zijn hetzelfde: geen significant verschil met Ellenberg ($p = 11,0\%$), terwijl het ecotopensysteem beter presteert ($p = 2,6\%$). De uitgebreide werkwijze die nodig is voor een analyse met het kiwa-systeem, leidt dus niet tot een verbetering. Mogelijk komt dit doordat slechts een klein gedeelte van de soorten gebruikt wordt.

Tabel 33: Vereenvoudigd voorbeeld van de uitvoer van het geautomatiseerde Kiwa-indicatorsoortensysteem voor een opname uit het landschapstype vennen, geassocieerd als Dopheide-verbond. Het Dopheide-verbond komt voor in de vochtclassen 2a – 5b. De klassen 2a – 4b worden door alle 7 soorten geïndiceerd. Deze klassen zijn dus zeer waarschijnlijk het milieu waarin de opname zich bevindt, en worden volledig geïndiceerd (code 'Zv'). Twee soorten indiceren daarnaast de klassen 5a en 5b, dus de waarschijnlijkheid dat deze het milieu van de opname representeren is laag (code 'L').

Soorten	CoverCode	Abundantie	min	2a	2b	3a	3b	4a	4b	5a	5b	max
417	6	0.18		p.i.	p.i.	p.i.	p.i.	p.i.	p.i.			
473	6	0.18		p.i.	p.i.	p.i.	p.i.	p.i.	p.i.	p.i.	p.i.	
777	2	0.02		p.i.	p.i.	p.i.	p.i.	p.i.	p.i.			
832	5	0.08		p.i.	p.i.	p.i.	p.i.	p.i.	p.i.	p.i.	p.i.	
849	2	0.02		p.i.	p.i.	p.i.	p.i.	p.i.	p.i.			
1069	2	0.02		p.i.	p.i.	p.i.	p.i.	p.i.	p.i.			
3016	5	0.08		p.i.	p.i.	p.i.	p.i.	p.i.	p.i.			
Som plekindicaties			0	7	7	7	7	7	7	2	2	0
Waarschijnlijkheid				Zv	Zv	Zv	Zv	Zv	Zv	L	L	

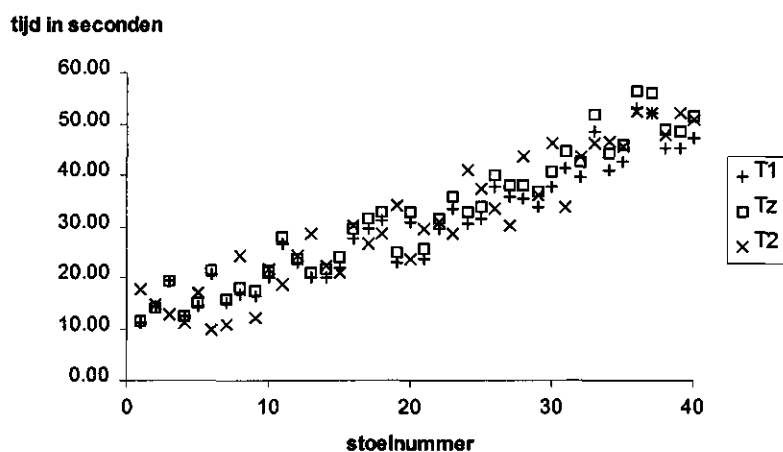


Figuur 33: De relatie van de GVG met de gemiddelde indicatiewaarden volgens Ellenberg en het ecotopensysteem (beide berekend met de kwalitatieve methode) en de indicatie volgens het Kiwa-indicatorsoortensysteem. De indicatiewaarden zijn gestandaardiseerd door van elke waarde het groepsgemiddelde af te trekken en te delen door de standaardafwijking, en ervoor te zorgen dat elke groep een positieve correlatie met de GVG heeft.

Bijlage 5 Voorbeeld permutatietest

Omdat enkele uitkomsten van de permutatietest vreemd over kunnen komen (bijvoorbeeld wanneer zeer kleine verschillen significant zijn en veel grotere verschillen niet), heb ik een eenvoudig experiment gesimuleerd (naar een idee van P. Torfs).

We hebben in een collegezaal een rij met 40 stoelen. We meten hoe lang een student er over doet om van zijn stoel de collegezaal uit te lopen. Daarvoor neemt deze student telkens op een van de 40 stoelen plaats, en loopt vervolgens naar buiten. De tijd die hij daarvoor nodig heeft (in seconden) wordt gesimuleerd door de functie $T = 10 + \text{stoelnummer} + \varepsilon_1$, waarin ε_1 (de 'fouteterm') uniform verdeeld is over het interval $[-7\frac{1}{2}, 7\frac{1}{2}]$. We laten deze student 2 keer op elke stoel zitten, zodat er uiteindelijk 2 meetwaarden voor elke stoel zijn (T_1 en T_2). Na de eerste keer hebben we ook gemeten hoe lang het duurt voordat de student vanaf zijn stoel de collegezaal uitgelopen is en het nummer van de stoel waar hij op zat kan zeggen (T_z). Naarmate de stoelnummers hoger worden, wordt dat moeilijker te onthouden en duurt het dus nog eens extra lang voor de student het stoelnummer kan zeggen. Deze tijd wordt gesimuleerd door $T_z = T_1 + \frac{\text{stoelnummer}}{10} + \varepsilon_2$. Hier is ε_2 uniform verdeeld over het interval $[-\frac{1}{2}, \frac{1}{2}]$. Voor alle 'meetwaarden', zie Figuur 34.

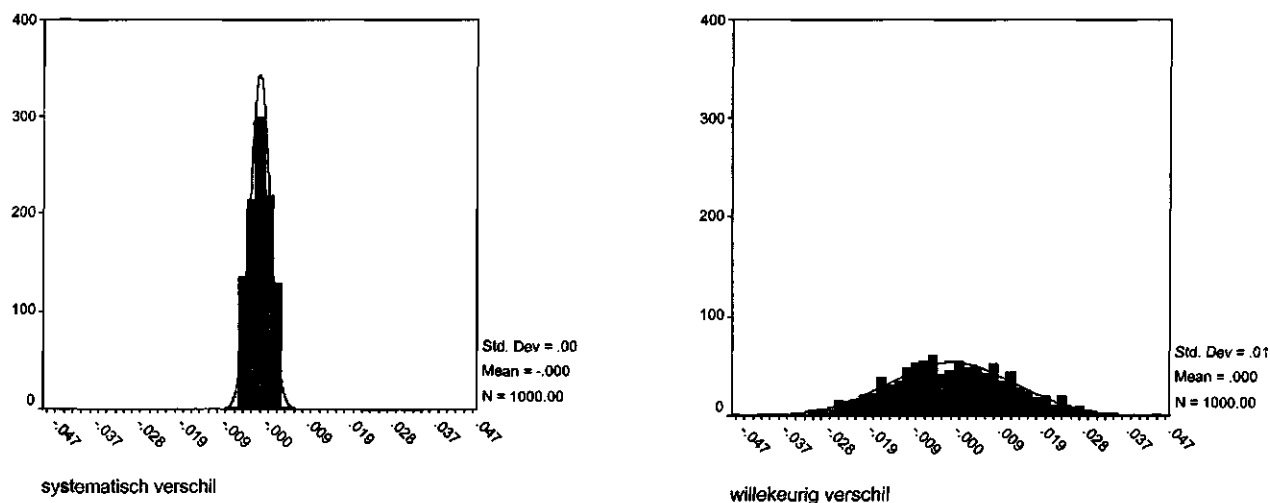


Figuur 34: De drie 'gemeten' tijden.

Zoals te zien is in Figuur 34, zijn de verschillen tussen T_1 en T_2 veel groter dan de verschillen tussen T_1 en T_z , terwijl de verschillen tussen T_1 en T_2 volledig op toeval berusten en de verschillen tussen T_1 en T_z grotendeels systematisch zijn. Voor de r_s geldt hetzelfde: de correlatie tussen het stoelnummer en T_1 is 0,965, tussen stoelnummer en T_z 0,970 en tussen stoelnummer en T_2 0,947. We gaan proberen uit te vinden met behulp van een permutatietest in hoeverre deze verschillen significant zijn.

De permutatietest zoals die beschreven is in paragraaf 2.4.3 geeft de verdelingen onder de nulhypothese 'geen verschil' (Figuur 35). De verdeling bij de test tussen T_1

en T_z is erg smal. Daardoor is het kleine verschil in r_s (0,005) significant ($p = 0,4\%$, tweezijdig). Het grotere verschil tussen T_1 en T_2 (0,018) is niet significant ($p = 19,2\%$, tweezijdig), vanwege de bredere verdeling onder de nulhypothese. Ter vergelijking: de test door middel van de z-transformatie van de Pearson correlatiecoëfficiënten, onder andere beschreven in Sachs (1984), geeft aan dat zowel het verschil tussen T_1 en T_z als tussen T_1 en T_2 niet significant is (resp. $p = 74\%$ en $p = 64\%$). Vooral het verschil tussen T_1 en T_z wordt door deze z-test zwaar onderschat.



Figuur 35: De verdelingen van de verschillen tussen twee correlatiecoëfficiënten in 1000 permutaties. Links de test voor verschillen tussen T_1 en T_z rechts tussen T_1 en T_2 .

Omdat hier gekozen is voor een eenvoudig te begrijpen voorbeeld, is deze test voor verschillen in de correlatiecoëfficiënt enigszins triviaal. We kunnen aannemen dat de tijd die nodig is om het stoelnummer te zeggen zelf ook positief gecorreleerd is met het stoelnummer: vandaar de hogere correlatie van T_z met het stoelnummer. Of de correlatie met T_z beter (in de kwalitatieve zin van het woord) is dan de correlatie met T_2 , is geen vraag van belang.

In deze studie naar methoden om indicatiewaarden voor vegetatieopnamen te berekenen, wordt aangenomen dat een significant hogere correlatiecoëfficiënt duidt op een systematisch betere methode. In die zin heeft een hogere correlatiecoëfficiënt wel een betekenis.