

# Plantengemeenschappen van Nederland

## Het gebruik van numerieke methoden bij het samenstellen van vegetatietabellen

S.M Hennekens

A.H.F. Stortelder



BIBLIOTHECA BOTANICA  
JOOP SCHAMINEE  
BOEKEN

R - 0700



Intern rapport  
Rijksinstituut voor Natuurbeheer  
Arnhem, Leersum en Texel

**PLANTENGEMEENSCHAPPEN VAN NEDERLAND**

**HET GEBRUIK VAN NUMERIEKE METHODEN BIJ HET SAMENSTELLEN  
VAN VEGETATIETABELLEN.**

S.M. Hennekens  
A.H.F. Stortelder

Tussentijdse rapportage van het project "Plantengemeenschappen van  
Nederland"

Intern rapport 90/18

Rijksinstituut voor Natuurbeheer  
Postbus 46  
3956 ZR Leersum

1990

Overname van informatie uit dit rapport is toegestaan onder voor-  
waarde van juiste referentie, met expliciete vermelding van de  
status.

F- 520243

## VOORWOORD

De tussentijdse resultaten van het project "Plantengemeenschappen van Nederland" zullen in de vorm van Interne RIN-rapporten als opeenvolgende delen worden uitgebracht. Uiteinderlijke publicatie zal, gebaseerd op rapportteksten (na het doorlopen van diverse commentaarronden), en aangevuld met illustraties, in boekvorm plaatsvinden. Gezien de omvang van het te schrijven boekwerk ligt het voor de hand het werk gefaseerd in afzonderlijke delen te laten verschijnen. Naast vier vegetatiebeschrijvende delen betreft dit ook een deel over gronslagen en methoden, inclusief handleiding. Het hier voor U liggende rapport zal als hoofdstuk in dit deel worden opgenomen.

Het project wordt uitgevoerd door het Rijksinstituut voor Natuurbeheer onder leiding van J.H.J. Schaminée. Voor wat betreft de syntaxonomische bewerking van de Nederlandse struweel- en bosgemeenschappen wordt samengewerkt met het Rijksinstituut voor Onderzoek in de Bos- en Landschapsbouw "De Dorschkamp (onder leiding van Dr. A.H.F. Stortelder).

Ter ondersteuning van het project is een drietal begeleidingscommissies ingesteld, te weten een Wetenschappelijke Begeleidingscommissie, een Gebruikerscommissie, en een Commissie Numerieke Verwerking. Deskundigen op deelgebieden, voor zover niet opgenomen in een van de genoemde commissies, krijgen de betreffende concepten ter commentaar toegestuurd. Verdere suggesties voor aanvulling en/of wijzigingen worden zeer op prijs gesteld. Deze kunnen worden toegezonden naar het op de titelpagina vermelde adres t.a.v. Drs. S.M. Hennekens.

Een groot aantal personen zijn we erkentelijk voor door hen verstrekte belangwekkende en opbouwende kritiek op het manuscript, dat als basis heeft gediend voor het onderhavige concepthoofdstuk. Met name willen we Jan Barkman, Rudi van Diggelen, Lodewijk van Duuren, Ladzi Fresco, Joop Schaminée, Karel Sykora en Onno van Tongeren bedanken.

De studie is mede mogelijk gemaakt door externe financiering van het Staatsbosbeheer, de Landinrichtingsdienst en het Natuur Milieu en Faunabeheer.

Foto omslag: Joop van Osch

## INHOUDSOPGAVE

Voorwoord

Inhoudsopgave

1. Inleiding .....	1
2. Transformatie, weging en standaardisering .....	3
3. Toelichting op de binnen "Plantengemeenschappen" gebruikte programmatuur .....	6
3.1 Inleiding .....	6
3.2 Het programma TWINSPAN .....	8
3.3 Indices en het programma FLEXCLUS .....	9
4. Gevolgde procedure bij de verwerking van vegetatie-opnamen in het project "Plantengemeenschappen van Nederland" .....	12
5. Enkele kanttekeningen bij het gebruik van numerieke methoden .....	14
6. Gerefereerde literatuur .....	17
7. Bijlagen .....	21
7.1. Bedekkingsschalen	
7.2 Procedureschema	

## HET GEBRUIK VAN NUMERIEKE METHODEN BIJ HET SAMENSTELLEN VAN VEGETATIETABELLEN

### 1. Inleiding

Gedurende de laatste decennia is een groot aantal computerprogramma's ontwikkeld voor het bewerken van vegetatie-opnamen. Het gaat hierbij met name om multivariate analyse-technieken. In de plantensociologie worden deze methoden gebruikt voor het vaststellen van numerieke betrekkingen tussen plantensociologische entiteiten (bijv. vegetatie-opnamen) en tussen deze entiteiten en milieufactoren. De multivariate analyse wordt in de Braun-Blanquet-methode toegepast in de synthetische fase van het onderzoek.

De functie van deze numerieke verwerkingsmethoden is tweeledig. Allereerst stellen de technieken de onderzoeker in staat om structuur in de gegevens vast te stellen en te visualiseren; op de tweede plaats worden de gegevens op een op reproduceerbare wijze samengevat, zijn gemakkelijk opnieuw te bewerken en goed overdraagbaar.

Numerieke methoden zijn al vroeg geïntroduceerd in de vegetatiekunde. Van der Maarel (1979a), die de geschiedenis van de numerieke classificatie ten behoeve van de syntaxonomie tot dan toe samenvat, noemt het werk van Sørensen (1948) als startpunt. Deze Scandinavische onderzoeker paste een similariteitsindex toe voor de vergelijking van vegetatie-opnamen en stelde een hiërarchie op van opnamengroepen, gebaseerd op floristische overeenkomsten tussen opnamen en opnamengroepen. Hij vergeleek de resultaten vervolgens met het systeem van Braun-Blanquet.

Na de opkomst van de computer werden de numerieke methoden op grote schaal toegepast. Het werd mogelijk om steeds meer opnamen steeds sneller te verwerken. Ofschoon de ontwikkeling van de methodologie met betrekking tot de numerieke verwerking van opnamen geleidelijk plaats vond, volgden de computerprogramma's elkaar snel op. Dit leidde tot een dermate groot aanbod van technieken, dat het nauwelijks mogelijk is om een goed overzicht te behouden. Samenvattende overzichten worden onder andere gegeven door Whittaker, 1967), Goodall (1970), Van der Maarel (1979a), Greig-Smith (1980), Jongman et al. (1987), Austin (1985), Noy-Meir & Van der Maarel (1987) en Kent & Ballard (1988), Mucina & Van der Maarel (1989).

Multivariate methoden worden gewoonlijk onderverdeeld in classificatie- en ordinatie-technieken. Zowel bij classificatie als bij ordinatie wordt de plaats die een plantensociologische entiteit krijgt toegewezen bepaald door de floristische samenstelling van de opnamen, waarbij bedekkingswaarden al dan niet een rol kunnen spelen (Van der Maarel, 1979a). Classificatie, in de plantensociologische context, kan omschreven worden als de procedure voor het onderschei-

den van vegetatietypen. Ordinatie is in dit verband de procedure van het rangschikken van opnamen en/of soorten in een één- of meerdimensionale ruimte. Meestal is de rangschikking gecorreleerd met de milieuvariatatie.

In de beginperiode van het gebruik van de multivariate technieken, werden classificatie- en ordinatiemethoden naast elkaar ontwikkeld. Er was veel discussie omtrent de betekenis en de keuze van beide typen van technieken (Kent & Ballard, 1988). Tegenwoordig is men van mening dat beide typen van methoden zeer goed in complementaire zin kunnen worden toegepast. Bij ordinatie ligt het accent op de rangschikking van de opnamen, terwijl bij classificatie de nadruk ligt op het onderscheiden van groepen van vegetatie-opnamen (zo homogeen mogelijke clusters, die onderling zoveel mogelijk verschillen).

Een van de problemen bij de toepassing van ordinatie- en classificatie-methoden is het kiezen van de juiste techniek uit een groot aantal mogelijkheden. Hoewel veel pogingen zijn gedaan om de toepasbaarheid van de afzonderlijke methoden ten aanzien van verschillende onderdelen te evalueren (zie o.a. Everitt (1974), Gauch (1982), Ter Braak (1978) en Van Tongeren (1987)), is hieruit evenwel nooit een eensluidende opvatting naar voren gekomen (zie o.a. Kent & Ballard, 1988; Mucina & Van der Maarel, 1989). De keuze voor een bepaalde methode wordt veelal niet gemotiveerd en heeft mede te maken met de beschikbaarheid van programma's. Voor de niet-specialist is het moeilijk om de optimale methode te kiezen uit het scala aan mogelijkheden. Het gevaar bestaat derhalve dat onvoldoende kennis van de gebruikte techniek kan leiden tot een onverantwoord, dat wil zeggen weinig kritisch gebruik.

In dit hoofdstuk wordt allereerst ingegaan op de kwantitatieve aspecten met betrekking tot de informatiewaarden van de soorten (par. 2). Er wordt aandacht besteed aan transformatie en standaardisatie van bedekkingswaarden en weging van soorten, gezien de grote invloed die deze kunnen hebben op de classificatie. In paragraaf 3 wordt een toelichting gegeven op de programmatuur die binnen het project "Plantengemeenschappen van Nederland" wordt toegepast. De gehanteerde procedure voor de verwerking van de vegetatie-opnamen tot tabellen wordt in paragraaf 4 behandeld (zie Fig. 1).

Met het toenemend gebruik van numerieke verwerkingsmethoden binnen de plantensociologie is ook de kritiek toegenomen op het handmatig samenstellen van vegetatietabellen, zoals dat voordien gebruikelijk was. Deze methode zou subjectief zijn, met andere woorden in hoge mate bepaald worden door de persoonlijke opvatting van de onderzoeker. Het toepassen van numerieke methoden daarentegen zou onbevooroordeeld zijn. De classificatie wordt in het laatste geval immers volgens tevoren opgestelde criteria uitgevoerd, om zodoende tot een objectieve syntaxonomische indeling te komen. Op deze discussie wordt in de laatste paragraaf nader ingegaan.

## 2. Transformatie, weging en standaardisering

Wanneer bedekkings- en/of abundantiewaarden betrokken worden bij de numerieke verwerking van vegetatie-opnamen, dient op de allereerste plaats de desbetreffende bedekkingsschaal getransformeerd te worden naar een rekenkundige schaal. Men denke hier bijvoorbeeld aan de codes r, +, 2m, 2a en 2b uit de door Barkman, Doing en Segal verfijnde schaal van Braun-Blanquet, die niet als zodanig aangewend kunnen worden voor numerieke verwerking. Van der Maarel (1979b) heeft deze schaal omgezet naar een schaal van 1 t/m 9.

Met betrekking tot de vegetatie-opnamen die binnen "Plantengemeenschappen" worden gebruikt is een viertal opname-schalen van belang. Dat is op de allereerste plaats de reeds genoemde schaal (door Barkman, Doing & Segal gemodificeerde schaal van Braun-Blanquet, inclusief de originele Braun-Blanquet-schaal), de schaal van Barkman, Doing en Segal (welke een andere voorstelt dan de eerstgenoemde), de schaal van Doing en als laatste, de zeer gedetailleerde schaal van Londo.

Andere schalen worden tot een van deze vier herleid. In bijlage 1 staan de vier schalen weergegeven, met daarbij de vertaling van de abundantie cq. bedekkingscode naar de procentuele waarde.

De naar procenten getransformeerde bedekkingswaarden kunnen, afhankelijk van de vraagstelling en de aard van het opname-materiaal nog transformaties en standaardisaties ondergaan.

Braun-Blanquet		Barkman, Doing & Segal		Van der Maarel
Symbool	Bedekking/ abundantie	Symbool	Bedekking/ abundantie	Symbool
		r	<< 5% enkele ind.	1
+	<5 % weinig ind.	+	< 5% weinig ind.	2
1	< 5% veel ind.	1	< 5% veel ind.	3
		2m	< %5 zeer veel ind.	4
2	5-25% of < 5%	2a	5-12,5%	5
		2b	12,5-25%	6
3	25-50%	3	25-50%	7
4	50-75%	4	50-75%	8
5	75-100%	5	75-100%	9

Figuur 1: Transformaties van de originele Braun-Blanquet-schaal.

Welk type transformatie het best toegepast kan worden hangt af van de aard van het opnamemateriaal. Een extreme vorm van transformatie is de "niet-transformatie"; hierbij ondergaan de bedekkingswaarden geen enkele verandering. Het is echter vrijwel uitgesloten dat een dergelijke transformatie tot een zinvol resultaat zal leiden, soorten met een relatief lage bedekking kunnen dan immers op geen enkele manier een rol spelen bij de indeling.

Een andere extreme vorm van transformatie is die waarbij bedekkingswaarden geen enkele rol meer spelen; er is alleen sprake van absentie of presentie van een soort. Een dergelijke transformatie kan zinvol zijn bij een heterogene set van opnamen. Bij een homogene set van opnamen zijn het vooral de bedekkingswaarden die aanleiding zijn tot het onderscheiden van verschillende opnamegroepen. Vooral bij soortenarme opnamen, zoals vaak het geval is bij dominantiegemeenschappen, is een extra accentuering van de bedekking van belang. In welke mate de bedekking benadrukt dient te worden, hangt ook af van de rang van de desbetreffende syntaxa. Jensen (1978) en Van der Maarel (1979b) stellen dat hoog in de hiërarchie van het classificatie-systeem (bijvoorbeeld op klasse- en orde-niveau) bedekkingen van soorten een ondergeschikte rol spelen en dat, naarmate men in de hiërarchie afdaalt kwantitatieve aspecten steeds meer van belang worden.

Eén van de meest toegepaste methoden voor de transformatie van bedekking/abundantie-cijfers is de hierboven genoemde Van der Maarel-transformatie (Van der Maarel, 1979b) (zie ook fig. 1). Hierin is de door Barkman, Doing en Segal (1964) gemodificeerde schaal van Braun-Blanquet omgezet in een 9-delige schaal (beide zijn ordinale schalen). Voor de bedekkingswaarden 5 t/m 9 (Braun-Blanquet-schaal: 2a t/m 5) komt dit enigszins overeen met een logaritmische transformatie; de gecombineerde schattingswaarden  $r$ ,  $+$  en  $1$  (Van der Maarel-schaal: 1 t/m 3) wijken daar natuurlijk vanaf, omdat hun logaritmen negatief of nul zijn. Deze benadering komt goed overeen met de oorspronkelijke opvatting van Braun-Blanquet ten aanzien van de gecombineerde schatting, waarbij relatief veel waarde wordt toegekend aan soorten, die weliswaar met een lage bedekking voorkomen, maar die wel met een relatief groot aantal individuen vertegenwoordigd zijn. In veel gevallen komen karakteristieke soorten van nature met een lage bedekking in de vegetatie voor. Juist hier speelt de abundantie een belangrijke rol bij de beoordeling of een dergelijke soort wel of niet optimaal in de opname voorkomt. Van Tongeren (1987) stelt dat een transformatie van reële bedekkingswaarden naar een ordinale schaal (dus een indeling in klassen) een verlies aan informatie betekent. Indien de reële voorhanden zijn verdient een ander soort transformatie de voorkeur, bijv. een logaritmische of een worteltransformatie.

Een intermediaire oplossing tussen het verwerken met absentie-presentie van soorten en het verwerken met inachtneming van reële bedekkingswaarden wordt gegeven door Hill et al. (1975), waarbij een soort door één of meerdere absentie/presentie-variabelen ("pseudo-soorten") vervangen wordt. Het principe is dat naarmate een soort een hogere bedekking heeft hij door meer pseudo-soorten vervangen wordt, waardoor de soort meer gewicht krijgt. In het programma TWINSpan, dat gebruik maakt van deze procedure, kan men door middel



van het definiëren van drempelwaarden het aantal pseudo-soorten dat toegewezen wordt bij een bepaalde bedekking zelf bepalen.

Het betrekken van de bedekkingswaarden van soorten in de numerieke verwerking van vegetatie-opnamen kan tot gevolg hebben dat dominante soorten een té grote rol gaan spelen ten opzichte van de overige soorten, die met lage bedekkingen voorkomen, maar wel een hoge diagnostische waarde kunnen hebben. Het inperken van de rol van hoge bedekkingen van soorten is ook om twee andere redenen hier op z'n plaats: ten eerste omdat dominantie vaak op storing wijst terwijl juist de achterliggende vegetatietypen van belang zijn en ten tweede omdat soorten vaak exponentieel groeien. Het kan daarom wenselijk zijn om de bedekkings- en abundantiewaarden eerst aan een transformatie te onderwerpen om op die manier dergelijke effecten te voorkomen. Het ligt dus ook om deze reden voor de hand om een logaritmische of een worteltransformatie toe te passen op de bedekkingswaarden.

Behalve door middel van transformatie kan ook door toekenning van een verschillend gewicht aan soorten de classificatie beïnvloed worden. Als men bijvoorbeeld de classificatie van bos-opnamen op grond van de spontane ondergroei wil laten plaatsvinden, dienen de soorten in struik- en boomlaag ondergewaardeerd te worden omdat deze (in Nederland) vaak zijn aangeplant of door selectieve kap sterk beïnvloed. Een lager gewicht wordt ook vaak toegekend aan "toevallige" soorten, indien de gebruikte (dis)similariteitsindex of de clustertechniek hiervoor zodanig gevoelig is, dat de aanwezigheid van deze min of meer toevallige soorten, leidt tot een sterke beïnvloeding van het resultaat (Van Tongeren, 1987). Ook voor wanneer soorten als systeemvreemd moeten worden opgevat (zie Schaminée et al., 1990) kan ook het lager waarderen van deze soorten tot betere resultaten leiden; dat wil zeggen de clusters zijn ecologisch beter te interpreteren.

Het toekennen van een hogere waardering aan soorten kan plaatsvinden, indien deze soorten blijkens de literatuur of blijkens de eerste resultaten van een classificatie-procedure of andere bewerking, een hoge diagnostische waarde hebben.

Standaardisering van soortenaantal of van bedekkingscijfers is een derde mogelijkheid om een groep van opnamen voorafgaand aan de clustering te bewerken. Standaardisering naar de maximale bedekking van soorten bijvoorbeeld vindt plaats omdat mag worden aangenomen dat kleine verschillen in bedekking bij soorten die nooit veel bedekken oecologisch even indicatief kunnen zijn als grote verschillen bij dominanten. Bij clustering van bos-opnamen bijvoorbeeld, wordt door standaardisering van de bedekking, het gewicht van een weinig abundant kruid gelijk aan dat van een dominante boomsoort.

Een andere, soms zinvolle standaardisering is die van de presentie van de soorten. Hierbij worden de bedekkingen van een soort over alle opnamen waarin deze voorkomt gesommeerd en gedeeld door het totaal. Hierdoor worden zeldzame soorten opgewaardeerd en wordt het gewicht van algemene soorten verminderd (Jongman et al., 1987). Het

is aan te bevelen een dergelijke standaardisering altijd gepaard te laten gaan met een onderwaardering van oecosysteem-vreemde soorten. De combinatie van een standaardisering van de bedekking en de presentie en van onderwaardering van de toevallige (vreemde) soorten benadert goed de klassieke tabellen-procedure volgens de opvattingen van de Braun-Blanquet school.

### **3. Toelichting op de binnen "Plantengemeenschappen" gebruikte programmatuur**

#### **3.1. Inleiding**

Voor het samenstellen van vegetatie-tabellen is binnen het project "Plantengemeenschappen" weinig direct gebruik gemaakt van ordinatietechnieken. Indirect echter veel, omdat bij het classificatieprogramma TWINSPAN, ordinatie ten grondslag ligt aan de indeling van de opnamen; daarnaast zijn ordinatietechnieken aanvullend gebruikt om meer inzicht te krijgen in de heterogeniteit binnen het materiaal. In de inleiding (Par. 1) zijn de begrippen ordinatie en classificatie reeds gedefinieerd. Hieronder wordt nader op deze begrippen ingegaan; het grootste gedeelte van deze paragraaf is gewijd aan classificatie, waarbij twee classificatie-programma's nader belicht worden.

Het begrip ordinatie is door Goodall (1954) geïntroduceerd en stamt van het Duitse Ordnung, een term die door Ramenski (1930) werd gebruikt om zijn directe ordinatiemethode te kwalificeren.

Ordinatie wordt door Goodall algemeen gedefinieerd als: een ordening of rangschikking van entiteiten in een één- of multidimensionale ruimte. In ons geval vertegenwoordigen de punten soorten, opnamen of tabellen.

Omdat het niet mogelijk is om de variatie in een twee-dimensionale figuur weer te geven wordt gebruik gemaakt van een aantal complementaire diagrammen. Elk ordinatie-diagram is een twee-dimensionale weergave, een projectie van alle punten in één vlak, met twee haaks op elkaar staande assen. Elk van deze assen verklaart een ander deel van de variatie. In de regel kunnen de belangrijkste assen, dat wil zeggen de assen die het grootste deel van de variatie weergeven, goed gecorreleerd worden met de meest relevante milieufactoren. Tegenover de indirecte ordinatie staat de directe ordinatie, waarbij de ordening van de punten (opnamen of soorten) is gebaseerd op het verband tussen milieufactoren en soorten. Voor een overzicht van ordinatie-technieken, zie Ter Braak (1987).

De meest gebruikte ordinatie-technieken zijn: PCA (Principal Component Analysis), CA (Correspondence Analyses, ook wel Reciprocal Averaging genoemd) en DCA (Detrended Correspondance Analysis), alle voorbeelden van indirecte ordinatie-technieken. RA is de techniek

die binnen TWINSpan gebruikt wordt voor rangschikking van opnamen en soorten.

Onder classificatie wordt verstaan: het toekennen van entiteiten (bijv. vegetatie-opnamen of taxa) aan groepen of klassen, op basis van hun onderlinge gelijkenis.

Ongeveer tot 1970 werden in de plantensociologie classificatiesystemen ontwikkeld op basis van met de hand gemaakte sociologische tabellen. Deze methode was zeer succesvol en leidde meestal tot dezelfde resultaten bij verschillende auteurs. Desondanks werd de methode met het de opkomst van computers steeds meer verlaten en vervangen door numerieke classificatiemethoden.

Bij de Braun-Blanquet-benadering worden voornamelijk hiërarchische cluster-methoden gebruikt. De vele voordelen hiervan zijn door Braun-Blanquet en zijn volgelingen vele malen uitvoerig uiteengezet.

De hiërarchische classificatiemethoden worden traditioneel onderverdeeld in twee typen:

1. agglomeratieve technieken, waarbij een serie van fusies plaatsvindt vanuit afzonderlijke objecten (opnamen) naar grotere groepen (clusters);
2. divisieve technieken, waarbij vanuit de totale verzameling van opnamen een serie splitsingen plaatsvindt naar steeds kleinere eenheden.

Agglomeratieve technieken beginnen met één enkele opname en voegen daar andere aan toe op grond van gelijkenis volgens bepaalde criteria, zodat een steeds grotere groep van opnamen ontstaat. Een nieuwe groep wordt pas gevormd wanneer geen enkele opname meer aan de voorwaarden voldoet om aan de bestaande groep(en) te worden toegevoegd. De criteria kunnen monothetisch of polythetisch van aard zijn. In het eerste geval vindt splitsing of fusie plaats op grond van aan- of aanwezigheid van slechts één soort, in het tweede geval is meer dan één soort bepalend voor de groepering van opnamen. Omdat in de Braun-Blanquet-school de classificatie wordt gebaseerd op de totale floristische samenstelling van de opnamen, zijn hierbij uiteraard uitsluitend polythetische criteria aan de orde.

Zowel agglomeratieve als divisieve technieken hebben voor- en nadelen. Voor divisieve technieken noemt Everitt (1974) o.a. het probleem van het niet kunnen herplaatsen (reallocceren) van opnamen die in een vroeg stadium zwak geassocieerd worden. Causton (1988) stelt dat een divisieve techniek superieur lijkt aan een agglomeratieve, omdat in het eerste geval wordt uitgegaan van de hele set van opnamen, die successievelijk in fragmenten wordt onderverdeeld, terwijl in het tweede geval het geheel vanuit een fragment opgebouwd moet worden. Dit geldt vooral, als uitgegaan wordt van een groot aantal opnamen. Bij bewerking van relatief kleine groep van opnamen daarentegen heeft een agglomeratieve techniek de voorkeur.

Van het grote aanbod van numerieke classificatietechnieken wordt voor plantensociologische doeleinden het programma TWINSpan (divisief) wel het meest gebruikt, zowel binnen als buiten Nederland. Gezien het feit dat ruime ervaring is opgedaan met dit programma en

dat het over het algemeen bevredigende resultaten oplevert, is ook voor het project "Plantengemeenschappen" hiervan veel gebruik gemaakt.

Daarnaast is ook het agglomeratieve programma FLEXCLUS toegepast. Bij grote groepen van opnamen is met behulp van TWINSPAN de indeling in clusters op een hoog hiërarchisch niveau opgespoord, bijv. voor de afbakening van de klassen, orden en verbonden. Met FLEXCLUS kunnen goede indelingen worden gemaakt wanneer gewerkt wordt met kleinere tabellen, zoals de groepen (clusters) van de lagere niveaus van grote TWINSPAN-bewerkingen.

### 3.2. Het programma TWINSPAN

TWINSPAN, Two-way Indicator Species Analyses (Hill et al, 1979) is een classificatieprogramma dat opnamen kan rangschikken aan de hand van differentiërende soorten, waarmee deze techniek enigszins vergelijkbaar is met de klassieke "handmethode". Het resultaat van het programma bestaat onder meer uit een gestructureerde tabel, waarbij de opnamen en de soorten zodanig geordend zijn dat in de matrix de getransformeerde bedekkingswaarden min of meer volgens een diagonaal zijn gerangschikt.

De eerste stap bij de TWINSPAN-bewerking is de ordinatie (Reciprocal Averaging) van de opnamen. De eerste ordinatie-as geeft de maximale spreiding van de opnamen weer, een ordening die overeenkomt met de belangrijkste floristische gradiënt. Iedere opname krijgt een score op de as en vervolgens wordt een (voorlopige) splitsing aangebracht op het zwaartepunt van de as. De opnamen aan de linkerkant van het nulpunt behoren tot de "negatieve" groep, die aan de rechterkant tot de "positieve" groep. Vervolgens wordt berekend hoe de verdeling van de soorten is in relatie tot hun voorkomen in de twee groepen. Omdat de eerste as de meeste variatie verklaart (en dus aangenomen mag worden dat ze de belangrijkste milieugradiënt voorstelt), kunnen de soorten die een sterke voorkeur hebben voor de ene dan wel de andere kant van het zwaartepunt van de as, voor de betreffende splitsing als differentiërende soorten beschouwd worden. Hill et al. (1975) spreken van "indicator species". Er moet op gewezen worden dat de begrippen "differentiërende soorten" en "indicator species" in dit verband een andere betekenis hebben dan de gangbare betekenis in de plantensociologie.

De eerste grove indeling wordt vervolgens verfijnd door een herberekening uit te voeren, waarbij uitsluitend de preferentiescores van de soorten gebruikt worden. Een absolute preferentiescore van 1 houdt in dat de presentie van een soort aan de ene kant van de dichotomy tenminste driemaal zo groot moet zijn dan aan de andere kant. Soorten met lage scores worden hierbij ondergewaardeerd. De verfijnde ordinatie (met kenmerken van CA en PCA) is een iteratief proces dat de uiteindelijke dichotomie bepaalt. Na deze tweede ordinatie rest nog de beoordeling van de zogenaamde "borderline cases", die dicht bij het punt liggen waarbij de verfijnde ordinatie

de set van opnamen is gedeeld. Voor deze gevallen wordt een derde ordinatie uitgevoerd, de zogenaamde indicator-ordinatie. Het doel van deze ordinatie is echter niet om de moeilijk te classificeren opnamen aan de ene of ander kant van de dichotomie te verplaatsen, maar om te bezien of op basis van een beperkt aantal hoog preferente soorten een "borderline"-opname van groep zou moeten veranderen. In het laatste geval wordt de opname door TWINSPAN als "misclassified" aangemerkt.

Na de definitieve splitsing kan de hele procedure opnieuw starten met elk van de twee gevormde groepen. Het maximaal aantal niveaus waarop op deze manier splitsingen uitgevoerd worden kan tevoren bepaald worden, evenals de minimale grootte van een opnamegroep (cluster).

Welke tweedelingen op een laag splitsingsniveau nog zinvol zijn is een vraag die bij de interpretatie moet worden beantwoord, al levert het programma hiervoor wel enige indicatie in de vorm van de zogenaamde eigenwaarde van de splitsing. In het algemeen geldt dat clusters die niet nader worden onderverdeeld en die dus overeenkomen met de syntaxa op het laagste hiërarchische niveaus laag in de boomstructuur worden aangetroffen. Ook het uiteindelijk aantal typen wordt derhalve bij interpretatie vastgesteld. Bij de beoordeling wordt aan verdere splitsingen van opnamengroepen geen nadere aandacht geschonken indien de floristische verschillen te gering zijn (twee clusters worden van elkaar onderscheiden door minstens één differentiërende soort).

Meer uitgebreidere beschrijvingen van het programma worden gegeven door Hill (1979) en Van Tongeren (1987).

### 3.2. Indices en het programma FLEXCLUS

Het principe van het interactieve programma FLEXCLUS (FLEXible CLUStering) (Van Tongeren, 1986) berust op de (dis)similariteit van opnamen: de mate van overeenkomst of van verschil tussen opnamen is bepalend voor de fusie. De similariteit of dissimilariteit tussen twee opnamen kan met behulp van bepaalde formules (indices) berekend worden.

In de loop der tijd is een groot aantal indices ontwikkeld, veelal toegesneden op een specifiek gebruik. Aan de basis van alle similariteits-indices ligt het idee ten grondslag dat de overeenkomst groter is naarmate twee opnamen meer soorten gemeenschappelijk hebben en in minder soorten van elkaar verschillen. Eén van de eerste indices was die van Jaccard (1912).

De Jaccard-index is als volgt gedefinieerd:

$$SJ=c/(a+b+c)$$

waarbij c het aantal overeenkomstige soorten voorstelt en a en b het aantal voor iedere opname unieke soorten.

Sørensen stelde in 1948 een enigzins andere similariteits-index voor, waarbij het aantal gemeenschappelijke soorten van beide opnamen niet wordt gedeeld door het totaal aantal soorten in beide opnamen, maar door het gemiddeld aantal soorten. De rol van de gemeenschappelijke soorten is hierbij dus van meer doorslaggevende betekenis, dan bij de Jaccard-index. De coëfficiënt van Sørensen wordt ook wel communitaets-coëfficiënt genoemd (CC). In formule:

$$SS (= CC) = 2c/(a+b+2c)$$

In geen van beide indices spelen presentie (bij vergelijking van tabellen) een rol. Kulczinski (1928) introduceerde overeenkomsten en verschillen in presentie in de formule van Jaccard. Ellenberg (1956) deed dit met de totale bedekkingswaarden der soorten. In het FLEXCLUS-programma wordt de similariteitsratio van Ball (1966) gebruikt. Deze index wordt bij de berekening van de similariteit tussen opname i en opname j voorgesteld door de volgende formule:

$$SR_{ij} = \sum_k y_{ki} y_{kj} / (\sum_k y_{ki}^2 + \sum_k y_{kj}^2 - \sum_k y_{ki} y_{kj})$$

Hierin is  $y_{ki}$  de abundantie van de k-de soort in opname i. Bij transformatie naar presentie-absentie, waarbij dus de verschillen in bedekkingswaarde wegvallen is deze formule identiek aan die van Jaccard.

Er zijn tal van andere indices voorgesteld voor het bepalen van de mate van overeenkomst tussen vegetatie-opnamen. De formule van Barkman (1958)

$$SR = C/\sqrt{ab}$$

geeft verschillen die zeer goed overeenkomen met wezenlijke verschillen, zoals die ook in de Braun-Blanquet-theorie onderkend worden. Beter is het

$$SR = C+1/\sqrt{ab} + 1$$

te gebruiken, omdat de noemer niet gelijk aan nul kan zijn. Het karakter van de index wordt hierdoor nauwelijks gewijzigd, behalve bij soortenarme opnamen.

Ondanks de vele studies naar de toepassingswaarde er van, bestaat volgens Kent & Ballard (1988) geen overeenstemming over welke index het best gebruikt kan worden onder welke omstandigheid. Volgens Van Tongeren (1987) wordt de keuze wordt vooral bepaald door de aard van het verzamelde materiaal, de ecologische vraagstelling en het persoonlijke oordeel van de onderzoeker.

Afhankelijk van bovengenoemde factoren zou in feite bij iedere nieuwe set te bewerken opnamen een bepaalde index gekozen moeten worden. Bijvoorbeeld in soortenrijk materiaal zou het accent meer moeten liggen op de kwalitatieve aspecten dan op de kwantitatieve en bij soortenarme gemeenschappen, waarin vaak dominante soorten optreden, geldt juist het omgekeerde. In het programma FLEXCLUS wordt gebruik gemaakt van een beperkt aantal indices, maar deze be-

perking wordt echter gecompenseerd door de mogelijkheid om, op tal van manieren, de rol van de bedekking van soorten te beïnvloeden (standaardisatie, transformatie en weging).

Het onderscheiden van groepen van opnamen wordt voorafgegaan door het berekenen van de overeenkomst op basis van een similariteitsmatrix (waarbij alle opnamen met elkaar worden vergeleken). Dit kan op verschillende manieren plaatsvinden. In FLEXCLUS wordt gebruikt gemaakt van een gewijzigde vorm van "single linkage clustering" toegepast, gevolgd door "centroid sorting" (cf. Gremmen, 1983). "Single linkage clustering" of "nearest neighbour clustering" houdt in dat opnamen worden samengevoegd tot kleine groepen, die op hun beurt weer aan elkaar worden gekoppeld, waarbij de afstand tussen twee groepen overeenkomt met de minimale afstand tussen een opname uit de ene groep en een opname uit de andere groep. Tevoren wordt echter een drempelwaarde gedefinieerd, voor de minimale gelijkheid die er tussen twee opnamen (of twee groepen van opnamen) vereist wordt, alvorens tot samengevoeging te komen (zie ook Sørensen, 1948). De methode van "single linkage clustering" werkt betrekkelijk snel, maar leidt, doordat vaak heterogene clusters ontstaan, meestal niet tot een bevredigend eindresultaat. Vandaar dat in het FLEXCLUS-programma deze procedure gevolgd wordt door een "centroid sorting procedure", waarbij de opnamegroepen als het ware opgeschoond worden. Voor elke opname wordt nagegaan of deze niet beter aan een andere groep kan worden toegevoegd. Een opname wordt verplaatst (gerelocerd) indien de afstand van die opname tot het centrum van de opnamegroep, waaraan deze in eerste instantie is toebedeeld, groter is dan de afstand die de opname heeft tot de centra van een van de andere groepen (Van Tongeren, 1987). Omdat na verplaatsing van een opname het centrum van zowel de groep waaruit de opname is verwijderd, als van de groep waaraan de opname is toegevoegd is veranderd, kan het zijn dat ook weer andere opnamen verplaatst dienen te worden, om de opnamegroepen maximaal van elkaar te differentiëren. Een aantal gevallen is het daarom zinvol om meer dan één herschikking plaats te laten vinden.

Na de herschikking kan nog op velerlei manieren het clusterresultaat bijgesteld worden. Zo kunnen sterk op elkaar lijkende groepen samengevoegd worden en toch nog te heterogene groepen worden opgesplitst. Ook kunnen opnamen uit een groep verwijderd worden, om op deze manier de homogeniteit van een cluster te vergroten. FLEXCLUS heeft verder als voordeel dat goed leesbare tabellen ontstaan. De volgorde van de opnamegroepen wordt bepaald aan de hand van hun positie op de eerste as van een ordinatie (reciprocal averaging). Deze volgorde kan interactief worden gewijzigd. Ook binnen de groepen worden de opnamen gerangschikt volgens een ordinatie. In de soortenlijst worden algemene soorten bovenaan de tabel geplaatst, daaronder volgen de differentiërende soorten en daaronder de indifferente soorten.

Behalve dat het programma FLEXCLUS zelf van individuele opnamen via berekening van similariteiten tot opnamegroepen kan komen bestaat er ook de belangrijke mogelijkheid de uitkomst van een TWINSPAN-bewerking te evalueren en desgewenst opnamen te verplaatsen en clusters te splitsen of te fuseren.

Een meer uitgebreide beschrijving van het programma FLEXCLUS wordt gegeven door Van Tongeren (1986).

#### 4. Gevolgde procedure bij de verwerking van vegetatie-opnamen in het project "Plantengemeenschappen van Nederland"

De vegetatie-opnamen, die de basis vormen van de indeling en de beschrijving van de Nederlandse plantengemeenschappen zijn afkomstig uit verschillende bronnen: veldboekjes (persoonlijke archieven), literatuur en geautomatiseerde archieven van diverse instituten en universiteiten.

Alle opnamen, die ter beschikking werden gesteld, dan wel uit literatuur en archieven werden achterhaald, werden eerst aan een kritische beschouwing onderworpen, alvorens ze al dan niet op te slaan in geautomatiseerde bestanden. De selectie-criteria die hierbij werden gehanteerd betreffen de volgende punten:

##### 1. Geografische spreiding.

Bij selectie van de opnamen is gestreefd naar een zo representatief mogelijke verzameling, d.w.z. dat:

- a. per locatie slechts een beperkt aantal opnamen van een syntaxon wordt ingevoerd (niet meer dan 5 tot 15);
- b. een groot deel van de gebieden waarin het syntaxon voorkomt vertegenwoordigd is met opnamen.

##### 2. Syntaxonomische spreiding.

Uitgaande van de oude indeling van Westhoff & Den Held (1969) is gestreefd naar een goede onderbouwing met opnamen van alle syntaxa van de klasse.

##### 3. Aard en kwaliteit van de opname.

- a. De taxa moeten op soortsniveau zijn onderzocht. Opnamen met veel aanduidingen op genus-niveau worden niet overgenomen (m.u.v. *Taraxacum*, *Rubus* en *Populus*), tenzij het slechts een enkele soort betreft met r of +.
- b. Voor die syntaxa waarvoor het relevant is dient de moslaag onderzocht te zijn.
- c. Er dient extra kritisch te worden gekeken naar publicaties (opnamen van auteurs) met een weinig professioneel karakter.

##### 4. Homogeniteit van de opname.

De opnamen dienen gemaakt te zijn volgens de beginselen van de Frans-Zwitserse school.

Wanneer echter het vermoeden bestaat dat zij betrekking hebben op heterogene proefvakken zijn ze voorzichtigheids-halve (althans in eerste instantie) uitgesloten. Dit vermoeden is soms uit de standplaatsbeschrijving (weiland met mieren hopen, helling 5 - 30 ), soms uit de soortensamenstelling af te leiden.

Ook opnamen gemaakt met de Tansley-schaal kunnen voor selectie in aanmerking komen indien men is uitgegaan van een homogene proefvlakte. De Tansley-methode stelt homogeniteit namelijk



niet als eis.

Voor de invoer van de opnamen in de computer zijn twee verschillende werkwijzen gevolgd:

- 1) via standaard-CBS/RIN-formulieren, door overschrijven vanuit opnameboekjes, en vervolgens 'vertoetsen' in een computerbestand (uitgevoerd door het CBS);
- 2) door opnamen één voor één rechtstreeks in te voeren in het computerbestand, met behulp van het interactieve programma TURBOVEG (Hennekens, 1990). Fouten worden hierbij geminimaliseerd, doordat tegelijkertijd controles worden uitgevoerd.

Behalve de floristische gegevens worden - zo mogelijk - ook de volgende, zogenaamde kopgegevens ingevoerd: volgnummer (is het unieke opnamennummer), geografische positie (kilometerhok of uurhok), datum van opname, bedekkingen en hoogten van boom-, struik- en kruidlaag, bedekking van de moslaag, oppervlakte opname, voorlopig syntaxon, auteur en opnameschaal. Volgnummer, hoknummer, syntaxon, auteur en opnameschaal zijn obligaat. Omdat een aantal opnameschalen geen directe rekenkundige verwerking toelaat wordt de abundantie- of bedekkingswaarde van een soort omgezet naar een procentuele schaal. Zowel de numerieke code als ook de oorspronkelijke code worden opgeslagen.

De schaal die bij het maken van opnamen meestal gehanteerd is, is de door Barkman gemodificeerde Braun-Blanquet-schaal; in mindere mate ook de Barkman, Doing & Segal-schaal, de Doing-schaal en de Londo-schaal. Wat betreft het syntaxon, dat bij elke opname vermeld wordt, geldt dat iedere opname een code krijgt toegewezen conform het syntaxonomische systeem van Westhoff & Den Held (1969). Na bewerking van een vegetatie-eenheid krijgen de betreffende opnamen een extra code volgens de nieuwe syntaxonomische indeling.

De invoer van de opnamen vindt plaats met behulp van het programma TURBOVEG, waarbij de kopgegevens en de floristische gegevens in aparte, maar wel aan elkaar gerelateerde bestanden zijn opgeslagen. Eenmaal opgeslagen in de computer kunnen naar diverse gezichtspunten selecties van opnamen plaatsvinden. Selectie is van belang omdat niet duizenden opnamen in een keer met behulp van classificatieprogramma's bewerkt kunnen worden.

Eén van de belangrijkste selectie-criteria is de syntaxonomische code. Omdat de revisie van de Nederlandse plantengemeenschappen in principe per klasse wordt aangepakt wordt meestal op klasse geselecteerd. Is er sprake van floristisch nauw verwante klassen (bijv. Molinio-Arrhenatheretea en Parvocaricetea) dan worden de opnamen uit deze klassen ook samen bewerkt, dat wil zeggen in één keer geclusterd. Verder is het ook mogelijk om een selectie te maken aan de hand van een tevoren geselecteerde, karakteristieke soortencombinatie. Ook het jaar van opname kan van belang zijn. Een selectie van opnamen van bijvoorbeeld vóór 1950 en ná 1950 maakt het mogelijk om verandering van de soortensamenstelling van de vegetatie in de tijd op te sporen.

Nadat een selectie is gemaakt, kunnen de opnamen uit het gegevensbestand gehaald worden, waarbij de soorten met de bedekkingen en de kopgegevens in aparte bestanden worden geplaatst. De soortsinformatie wordt per opname volgens een bepaald (internationaal) formaat, het Cornell-Condensed-format, gerangschikt. Dit formaat maakt bewerking met diverse ordinatie- en classificatie-programma's mogelijk.

Het Cornell-condensed-bestand wordt allereerst bewerkt met behulp van het programma CEDIT (Van Tongeren, 1990). Met dit programma worden bedekkingswaarden van de soorten onderworpen worden aan een transformatie. Verder kunnen nauwverwante taxa samengevoegd worden tot één taxon, bijvoorbeeld in geval dat *Eleocharis palustris*, *E. palustris* ssp. *palustris* en *E. palustris* ssp. *uniglumis* in dezelfde tabel voorkomen.

Vervolgens wordt met het al dan niet gemodificeerde bestand een TWINSPAN-classificatie uitgevoerd. Het resultaat hiervan kan, als een nadere classificatie-analyse is gewenst, opgesplitst worden in kleinere deelbestanden, die op hun beurt verder met TWINSPAN en/of FLEXCLUS bewerkt kunnen worden. Bij grote bestanden dient men er rekening mee te houden dat het aantal opnamen dat op een Personal Computer verwerkt kan worden nogal beperkt is. Vaak zal voor bewerking van een groot aantal opnamen toevlucht moeten worden genomen tot een computer met een groter werkgeheugen. De grootte van het vereiste werkgeheugen hangt samen met het aantal soorten, het aantal drempelwaarden voor de bedekking en het aantal opnamen.

Voor een goede interpretatie van de tabellen is het van belang dat de zogenaamde kopgegevens (auteur, datum, bedekkingen en hoogten van de verschillende vegetie-lagen, aantal soorten etc.) direct te lezen zijn. Daartoe kunnen met behulp van het programma TABELKOP (Hennekens, 1990) de kopgegevens boven een TWINSPAN- en FLEXCLUS-tabel geplaatst worden, hierbij worden tevens de moeilijk leesbare soortscodes vervangen door volledige soortsnamen.

De volgende stap in de procedure is de omzetting van de plantensociologische tabel, een matrix van vegetatie-opnamen tegen soorten) in een synoptische tabel (een matrix van opnamegroepen tegen soorten). Gegeven de groepsindeling van de opnamen wordt met behulp van het programma CEDIT (Van Tongeren, 1990) voor iedere soort per cluster de presentie- en karakteristieke bedekkingswaarde berekend. Onder karakteristieke bedekkingswaarde wordt verstaan de gemiddelde bedekking gedeeld door de presentie.

## 5. Enkele kanttekeningen bij het gebruik van numerieke methoden

In het algemeen is men het erover eens dat voor de verwerking van honderden of zelfs duizenden opnamen numerieke classificatiemethoden wenselijk zijn, zo niet noodzakelijk. Met behulp van een computer is men in staat om grote groepen van opnamen op te delen in kleinere groepen die ieder apart wél zonder hulp van computerprogramma's te overzien, d.w.z. te classificeren, zijn. De vraag is echter hoe de computer optimaal te hanteren en aan de hand van welke

(objectieve) criteria het resultaat van een numerieke bewerking beoordeeld kan worden.

Wat betreft het resultaat van een numerieke bewerking kan gesteld worden dat een classificatie-systeem pas waarde heeft als het bruikbaar is, met andere woorden: ecologisch goed interpreteerbaar. De totstandkoming van een opname is geen objectieve aangelegenheid. De keuze van de ligging en de grootte van de opname (onder meer bepalend voor de homogeniteit van de proefvlakte), en vooral ook de beslissing waar de opname gemaakt wordt, hangen sterk af van de ervaring en het persoonlijke oordeel van de onderzoeker. Met betrekking tot het gebruik van opnamen uit de literatuur dient er op gewezen te worden dat verschillende onderzoekers ieder min of meer hun eigen methode volgen en hun eigen (al dan niet systematische) fouten maken. Voorzover dit overzien kan worden kan hiermee echter bij de beoordeling van de tabellen rekening worden gehouden. Een ander probleem dat zich voordoet bij de beoordeling van de steekproef is de geografische verdeling van de opnamen en de verdeling over de syntaxonomische eenheden. Een kritische houding bij elke stap in het proces van classificatie (van opname tot eindtabel) blijft derhalve van het grootste belang.

Numerieke verwerking van opnamen kan niet benaderd worden met de "trechter-methode". Hiermee wordt bedoeld dat alle opnamen boven in een trechter (computer) worden gestopt en aan de onderkant keurig geclassificeerd eruit komen, waarbij de computer een soort "black-box" is waarbinnen het wonder gebeurt. Uit het voorgaande moge blijken dat dit echter niet het geval kan zijn en dat niet een bepaalde verwerkingsmethode toereikend is. In de loop der tijd zijn diverse methoden ontwikkeld, die alle toegesneden zijn op een specifiek gebruik. Anders gezegd, de aard van het opnamen-materiaal en de vraagstelling van de onderzoeker bepalen in hoge mate de te gebruiken methode.

Eén van de problemen, die samenhangen met de numerieke verwerking van opnamen, is het bezwaar dat uitsluitend de floristische samenstelling (al dan niet met inbegrip van de bedekkingswaarden) in de berekeningen wordt betrokken, terwijl bij de classificatie volgens de Braun-Blanquet-school ook met een ander vegetatiekundig aspect rekening wordt gehouden, namelijk de vegetatiestructuur.

Een ander probleem vormt het verschil in de ecologische informatiewaarde van soorten. In de Braun-Blanquet-school wordt aan soorten met een nauwe ecologische amplitudo, zoals bijv. bij orchideeën vaak het geval is (al dan niet onbewust), een groter gewicht toegekend. Volgens Barkman zou voor een objectieve benadering van het classificatie-vraagstuk aan iedere soort een bepaald gewicht toegekend moeten worden, dat overeenkomt met zijn informatiewaarde (die dan alleen regionaal geldig is).

Naast de verschillen tussen soorten kunnen verschillen binnen de soort ten aanzien van vitaliteit een rol spelen als het gaat om het onderkennen van vegetatie-eenheden. Populaties van *Orchis purpurea* bijvoorbeeld, kunnen bloeiend dan wel vegetatief voorkomen, hetgeen samenhangt met een verschil in standplaats. Hetzelfde geldt voor bloeiende Waterlelie's, die een andere diagnostische betekenis hebben dan niet bloeiende; de laatste bevinden zich in een later

stadium van de successie. In de beschikbare programma's kan deze informatie echter niet rekenkundig verwerkt worden.

Volgens welke criteria opnamen-groepen van elkaar dienen te worden gescheiden, en wat de syntaxonomische status is van het onderscheid (niveau van klasse, orde, verbond, associatie of subassociatie) hangt onder meer af van het totale aantal soorten van de opnamen. Een verschil van twee soorten voor opnamen van Lemna-begroeiingen kan aanleiding zijn tot het onderscheiden van twee verbonden, terwijl een verschil van twee soorten voor opnamen van bossen nog niet aanleiding hoeft te zijn tot het onderscheiden van twee aparte subassociaties.

Voor het beoordelen van de status van de hogere eenheden (m.n. op klasse- en orde-niveau) kan niet uitsluitend worden uitgegaan van Nederlands opnamen-materiaal. Hiervoor zullen mede de opvattingen in de ons omringende landen verkend moeten worden (beoordelen van buitenlandse opnamen in relatie tot Nederlandse en interpretaties in de buitenlandse literatuur). Een voorbeeld waarbij buitenlandse opvattingen doorslaggevend zullen zijn is de syntaxonomische positie van de Nederlandse muurvegetaties, onder te brengen in één dan wel twee klassen (*Parietarietea judaicae* en *Asplenietea rupestris*).

Uit het bovenstaande moge blijken dat de informatie, die in de opname besloten ligt, heeft niet alleen betrekking op de floristische samenstelling en de bedekking van soorten. Ieder begroeiings-type behoeft een eigen aanpak. Aan de hand van de ecologische interpreteerbaarheid van het resultaat zal moeten blijken of de gevolgde methode de juiste was. Van een zuiver objectieve classificatie kan derhalve geen sprake zijn, hoewel de rekenmethoden op zich zelf wel objectief zijn. De keuzes die echter gemaakt worden ten aanzien van het te hanteren computer-programma en vervolgens de interactieve beslissingen binnen het programma, zijn in zekere zin subjectief.

Door het toenemend gebruik van de computer bestaat het gevaar dat het gebruik ervan te veel wordt gezien als een doel op zich en veel minder als een middel om vegetatie-opnamen te bewerken (zie ook Kent & Ballard, 1988). Er dreigt daarbij een vervreemding op te treden van het object, de planten en de vegetatie. Vahle & Dettmar (1988) spreken van toenemende digitalisering van de vegetatie. Diverse auteurs wijzen op de relatieve betekenis van computerbewerkingen. Wiegleb (1986) merkt in dit kader op dat de vegetatie meer is dan een data-matrix en dat de onderzoeker in het veld, die de vegetatie-opnamen maakt, meer is dan een data-knecht. Debus (1985) maakt onderscheid tussen echte en kunstmatige informatie. De echte informatie komt voort uit een begrijpend bewustzijn en heeft een geestelijke oorsprong. Deze kennis kan men begrijpen, doorgronden en men kan teruggaan tot de geestelijke kern en de bedoeling ervan. Kunstmatige informatie ontstaat doordat eerst echte informatie wordt geformaliseerd, dan in de computer wordt ingevoerd en vervolgens wordt omgevormd tot nieuwe geformaliseerde informatie. Deze laatste vorm van informatie is niet meer gericht op de feitelijke fenomenen,

maar op de afgeleide hiervan, hetgeen zelfs kan leiden tot verkeerde conclusies. De uitspraak van Tüxen "Wer viel rechnet, muss damit rechnen, dass er sich ver-rechnet" (in Dettmar & Vahle, 1988) geeft op treffende wijze de kern van dit probleem weer, wellicht meer dan Tüxen zal hebben bedoeld.

Vergelijkbaar met het onderscheid tussen kunstmatige en echte informatie maakt Kleefman (1980) een onderscheid tussen eerste-orde en tweede-orde kennis. Onder het eerste verstaat hij de doorleefde kennis van het object, terwijl het tweede, de kwalitatief veel lager te waarderen kennis is die welke wordt verkregen aan de hand van een beperkt aantal waarnemingen en manipulaties hiervan. Ook in de plantensociologische context leek het ons zinvol om hier in het kort te wijzen op het gevaar dat de voor het object essentiële kenmerken en relaties tussen die kenmerken door een consequent doorgevoerde parametrische benadering over het hoofd worden gezien. Met dit voor ogen biedt de numerieke verwerking een scale van mogelijkheden voor een snelle en systematische verwerking van grote archieven.

## 6. Gerefereerde literatuur

- Austin, M.P. (1985). Continuum concept, ordination methods and niche theory. *Ann. Rev. Ecol. Syst.* 16: 19-31.
- Barkman, J.J. (1989). Fidelity and character-species, a critical evaluation. *Vegetatio* 85(1/2): 105-116.
- Barkman, J.J., H. Doing & S. Segal (1964). Kritische Bemerkungen und Vorschläge zur quantitativen Vegetationsanalyse. *Acta Bot. Neerl.* 13: 394-419.
- Ball, G.H. (1966). A comparison of some cluster seeking techniques. Stanford Research Institute, California.
- Braak, C.J.F. ter (1987). Ordination. In: Jongman et al. *Data analysis in community and landscape ecology*. Pudoc, Wageningen.
- Causton, D.R. (1988). *Introduction to vegetation analysis*. Unwin Hyman, London. 342 p.
- Debus, M. (1985). *Der Mensch im "Zugriff" des Computers*. Stuttgart.
- Everitt, B.S. (1974). *Cluster analyses*. Heinemann, London. 122 p.
- Gauch, H.G. (1982). *Multivariate analysis in community ecology*. Cambridge University Press, Cambridge. 298 p.
- Goodall, D.W. (1954). Objective methods for the classification of vegetation. *Australian Journal of Botany* 1: 39-63.
- Goodall, D.W. (1970). Statistical Plant Ecology. *Annual review of Ecology and Systematics*: 90-124.
- Greig-Smith, P. (1980). The development of numerical classification and ordination. *Vegetatio* 42: 1-9.
- Gremmen, N.J.M. (1983). Numerieke verwerking van vegetatiekundige gegevens. *Werkverslag Rijksinstituut voor Natuurbeheer*. 32 p.

- Hill, M.O. (1973). Reciprocal Averaging: an eigenvector method of ordination. *J. Ecol.* 61: 237-249.
- Hill, M.O. (1979). TWINSpan - A FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes. Cornell University Ithaca, New York. 90 p.
- Hill, M.O., R.G.H. Bunce & M.W. Shaw (1975). Indicator species analyses, a divisive polythetic method of classification and its application to a survey of native pinewoods in Scotland. *J. Ecol.* 63: 597-613.
- Jaccard, P. (1912). The distribution of the flora of the alpine zone. *New Phytologist* 11: 37-50.
- Jensen, S. (1978). Influences of cover values on classifications and ordination of lake vegetation. *Vegetatio* 37 (1): 19-31.
- Jongman, R.H.G., C.J.F. ter Braak & O.F.R. van Tongeren (1987). Data analysis in community and landscape ecology. Pudoc, Wageningen. 299 p.
- Kent & Ballard (1988). Trends and problems in the application of classification and ordination methods in plant ecology. *Vegetatio* 78(3): 109-125.
- Kleefman, F. (1980). Theoretische achtergronden van onderzoeksmethoden in de ruimtelijke ordening. *Landbouwkundig Tijdschrift* 92(48):102-111.
- Maarel, E. van der (1979a). Multivariate methods in phytosociology, with reference to the Netherlands. In: M.J.A. Werger (red.) "The study of vegetation". Junk, The Hague: 161-225.
- Maarel, E. van der (1979b). Transformation of cover-abundance values in phytosociology and its effects on community similarity. *Vegetatio* 39: 97-114.
- Mucina, L. & E. van der Maarel (1989). Twenty years of numerical syntaxonomy. *Vegetatio* 81: 1-15.
- Noy-Meir, I. & E. van der Maarel (1987). Relations between community theory and community analysis in vegetation science: some perspectives. *Vegetatio* 69: 5-15.
- Ramenski, L.G. (1934). Zur Methodik der vergleichenden Bearbeitung und Ordnung von Pflanzenlisten und anderen Objekten, die durch mehrere, verschiedenartig wirkenden Faktoren bestimmt werden. *Beiträge zur Biologie der Pflanzen* 18: 269-304.
- Schaminée, J.H.J., A.H.F. Stortelder & V. Westhoff (1990). Plantengemeenschappen van Nederland. De identificatie en classificatie van plantensociologisch onverzadigde gemeenschappen. Intern Rapport, Rijksinstituut voor Natuurbeheer. 19 p.
- Sørensen, T. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species content. *Det. Kong. Danske Vidensk. Selsk. Biol. Skr. (Copenhagen)* 5(4): 1-34.
- Tongeren, O. van (1986). FLEXCLUS, an interactive program for classification and tabulation of ecological data. *Acta Bot. Neerl.* 35(3): 137-142.
- Tongeren, O. van (1987). Cluster analyses. In: Jongman et al. Data analysis in community and landscape ecology. Pudoc,

Wageningen.

- Tongeren, O. van (1990). CEDIT. Manual. Limnological Institute. In prep.
- Vahle, H.-C. & J. Dettmar (1988). "Anschauende Urteilskraft", - ein Vorschlag für eine Alternative zur Digitalisierung der Vegetationskunde. *Tuexenia* 8: 407-415.
- Westhoff, V. & A.J. den Held (1969). *Plantengemeenschappen in Nederland*. Thieme, Zutphen. 324 p.
- Wiegand, G. (1986). Grenzen und Möglichkeiten der Datenanalyse in der Pflanzenökologie. *Tuexenia* 6: 365-377.
- Whittaker, R.H. (1967). Gradient analysis of vegetation. *Biological Reviews* 49: 207-264.
- Whittaker, R.H. (1973). Approaches of classifying vegetation. In: Whittaker R.H. (Editor): *Ordination and classification of communities*. Junk, The Hague: 325-354.

**BIJLAGEN**



## 7.1. Bedekkingsschalen

### Schaal 1: Braun-Blanquet (getransformeerd)

	Code	%
1	r	1
2	+	1
3	1	2
4	2m	4
	2	13
	2a	8
	2b	18
	3	35
	4	60
	5	85

### Schaal 2: Doing

	Code	%
	r	1
	p	1
	a	2
	m	4
	01	10
	02	20
	03	30
	04	40
	05	50
	06	60
	07	70
	08	80
	09	90
	10	97

### Schaal 3: Barkman, Doing & Segal

	Code	%
	r	1
	+r	1
	+p	1
	+a	1
	+b	2
	1p	1
	1a	2
	1b	3
	2m	4
	2a	8
	2b	18
	3a	31
	3b	43
	4a	56
	4b	68
	5a	81
	5b	93

### Schaal 4: Londo

	Code	%
	r1	1
	r2	2
	r4	4
	p1	1
	p2	2
	p4	4
	a1	1
	a2	2
	a4	4
	m1	1
	m2	2
	m4	4
	1	10
	1-	7
	1+	12
	2	20
	3	30
	4	40
	5	50
	5-	47
	5+	52
	6	60
	7	70
	8	80
	9	90
	10	97

Kolom 1: Oorspronkelijke code voor abundantie/bedekkingswaarde.  
 Kolom 2: Naar percentages getransformeerde schaal (1-99).