

Het identificeren van vegetatietypen met behulp van de computer

R. Pot

Identificatie is het beoordelen tot welk type of syntaxon uit een gepubliceerd syntaxonomisch overzicht een willekeurige vegetatie kan worden gerekend. De te identificeren vegetatie kan een enkele opname of een reeks van opnamen betreffen, maar ook een type uit een lokale typologie. In deel 1 van 'De Vegetatie van Nederland' is een hoofdstuk speciaal gewijd aan syntaxonomische identificatie (Schaminée et al. 1995b). Verschillende methoden worden beschreven, maar een duidelijke instructie ontbreekt. Bij nadere beschouwing is dat niet echt verwonderlijk, want hoewel vegetatiekundigen met kennis en ervaring weliswaar in staat zijn syntaxa te herkennen, is het erg moeilijk dit herkenningsproces formeel te beschrijven. Om die reden is de ontwikkeling van computerprogramma's voor identificatie van plantengemeenschappen, in tegenstelling tot de situatie bij het synthetisch onderzoek (clustering en ordinatie), nog nauwelijks van de grond gekomen. Aanzetten ertoe zijn onder andere gemaakt door Hill (1993) en Van Tongeren (ongepubliceerd).

Dit artikel beschrijft een benadering die gebaseerd is op de deductieve methode, zoals die is beschreven door Kopecký & Hejný (1974). Kopecký et al. (1995) publiceerden hiervoor een algoritme (een reeks logische rekenregels, te gebruiken als basis voor een computerprogramma) dat hier nader is uitgewerkt en is toegepast in het programma SynDiaT (Pot 1997).

Twee benaderingswijzen

In principe zijn er twee benaderingswijzen voor identificatie van syntaxa. De meest logische wijze lijkt vergelijking van de totale soortensamenstelling met die van alle typen. Veel efficiënter echter, en van oudsher veel meer toegepast, is het identificeren aan de hand van diagnostische soorten. Met 'Plantengemeenschappen in Nederland' (Westhoff & Den Held 1969) was het overigens ook niet mogelijk de totale soortensamenstelling bij de diagnose van een plantengemeenschap te betrekken: er werden alléén diagnostische soorten vermeld. Dit wordt anders met het verschijnen van 'De Vegetatie van Nederland' waarin samenvattende (synoptische) tabellen met alle relevante soorten worden gepresenteerd (Schaminée et al. 1995a, 1995c, 1996, in prep.).

Het is verleidelijk om een computerprogramma te ontwikkelen dat vergelijkingen maakt tussen willekeurige opnamen en de samenvattende tabellen. De daarvoor benodigde techniek is reeds goed ontwikkeld voor de synthetische fase van het vegetatieonderzoek en is gebaseerd op het gebruik van een similariteits-index.

Ook theoretisch lijkt dit goed verdedigbaar: wanneer een te identificeren opname door zo'n programma wordt geplaatst in een type, dan wordt die opname op dezelfde manier geordend als de opnamen die dat type oorspronkelijk beschreven en behoort dus tot dat zelfde type. Problematisch wordt

het alleen als de opname ergens tussen twee of meer typen in wordt geplaatst.

Er doen zich echter nog meer problemen voor, die van zeer grote praktische betekenis zijn. In de meeste opnamen hebben algemene soorten met een brede oecologische amplitudo numeriek de overhand. Daardoor lijkt zo'n opname op veel typen een beetje en is het heel lastig te bepalen op welk type hij het meest lijkt.

Specifiekere soorten zijn meestal in de minderheid en worden daardoor bij evenredige vergelijking doorgaans lager gewaardeerd, hoewel ze oecologisch gezien wel de meeste indicatiewaarde hebben.

Ter compensatie hiervan zouden de specifieke soorten kunnen worden opgewaardeerd, maar om dat objectief te doen, moet een computerprogramma de mate van specificiteit van de soort kunnen bepalen. Dit kan bijvoorbeeld door de trouwgraad van elke soort uit te rekenen per type, maar in feite is het programma dan iedere keer aan het berekenen welke soorten diagnostisch zijn.

Bij elke typologie worden soorten onderscheiden die kenmerkend zijn voor bepaalde typen of groepen van typen en daardoor diagnostisch ten aanzien van de identificatie van die typen.

In de syntaxonomisch indeling volgens Braun-Blanquet worden deze soorten kensoorten, differentiërende soorten en constante soorten genoemd, al naar gelang de reikwijdte van hun onderscheidend vermogen. Vaak worden ook nog verfijningen gebruikt, waarover later. De typen zijn daarbij ook nog hiërarchisch geordend. De associatie wordt daarbij beschouwd als de eenheid die het dichtst staat bij de concrete plantengemeenschap. De daarboven geordende verbonden worden beschreven door de voor groepen van associaties kenmerkende soorten, de diagnostische soorten op dat niveau. Daarboven geldt hetzelfde voor de orden en daarboven voor de klassen.

In een goed ontwikkelde ('verzadigde') plantengemeenschap komt een substantieel deel van de diagnostische soorten voor van zowel de associatie waartoe de vegetatie kan worden gerekend als van de erboven geordende niveaus. Samen met (een doorgaans groter aantal) niet-diagnostische soorten vormen deze soorten samen de totale soortensamenstelling. De niet-diagnostische soorten kunnen met een zekere regelmaat worden aangetroffen in een plantengemeenschap die kan worden gerekend tot die associatie, maar minstens even vaak of vaker ook in andere.

Aan het gebruik van diagnostische soorten kleeft vooral het bezwaar dat in een willekeurige opname soorten voorkomen die voor een scala aan syntaxa een diagnostische waarde blijken te bezitten. Vaak is het moeilijk vast te stellen welke soorten dan het zwaarst wegen.

Om dat laatste op te lossen is het nodig om inzicht te hebben in de onderlinge verhoudingen van de syntaxa; men moet kunnen bepalen wanneer men voldoende diagnostische soorten heeft om van een goed ontwikkeld syntaxon te kunnen spreken en men moet weten in hoeverre diagnostische soorten van het ene syntaxon als begeleidende soorten bij het andere syntaxon voorkomen. Dit vereist een gedegen kennis en veel ervaring.

De Braun-Blanquet-methode leidt langs een inductieve weg tot de opstelling van een hiërarchische indeling, dat wil zeggen door samenvoeging van typen tot typen van een hoger niveau. Gestructureerde identificatie vindt juist andersom plaats, langs deductieve weg. Eerst wordt met behulp van de diagnostische soorten voor de klassen gekken tot welke klasse de opname kan worden gerekend. Vervolgens wordt hetzelfde gedaan op elk lager syntaxonomisch niveau binnen het geïdentificeerde syntaxon. Gedeeltelijk wordt hier al rekening gehouden met de totale soortensamenstelling door de

kensoorten van de lagere syntaxonomische niveaus ook als diagnostisch te zien voor de hogere niveaus.

Deze werkwijze heeft een zekere analogie met de taxonomie. Ook hier worden soorten op grond van overeenkomstige kenmerken samengevoegd tot genera en verder, en soorten gedetermineerd door te bepalen met de kenmerken van de hoogste taxonomische niveaus.

Computerprogramma

De beschouwingen hierboven leiden mijns inziens tot de conclusie dat identificatie van typen in een hiërarchisch geordend overzicht waarbij diagnostische soorten zijn gegeven van elk type, in eerste instantie het beste kan geschieden op basis van de deductieve methode. In de vele gevallen echter waarin meer dan één type voor identificatie in aanmerking komt, lijkt het zinvol om de totale soortensamenstelling te gebruiken bij de afweging tussen de mogelijke oplossingen.

Het maken van een computerprogramma dat deze weg volgt is vrij ingewikkeld en ook het gebruik ervan is niet zonder haken en ogen. Op vele momenten moeten afwegingen worden ingebouwd die van cruciale invloed kunnen zijn op het resultaat.

Een algoritme voor zo'n programma is gepubliceerd door Kopecký et al. (1995). Dit algoritme heeft nogal wat beperkingen en werkt alleen redelijk goed in 'ideale' situaties, waarover later. Met dit algoritme als uitgangspunt is het computerprogramma SynDiaT gemaakt, een acroniem van Syntaxonomische Diagnostiek met Tabellen, of Syntaxonomical Diagnostics Tool.

Bij de analyse wordt eerst van alle syntaxa de diagnostische soorten (gewogen) geteld. De kensoorten worden daarbij ook als diagnostisch beschouwd bij alle hiërarchisch

hogere syntaxa. Daarna wordt vastgesteld welke klasse de hoogste score bereikt, daarbinnen welke orde de hoogste score bereikt, enzovoorts, tot op het niveau van (sub-)associatieniveau. In het ideale geval blijken alle soorten van de opname vervolgens ofwel diagnostisch op een van de hiërarchische niveaus van deze (sub-)associatie, ofwel als begeleidende soorten beschreven van de (sub-)associatie.

De weging van de diagnostische soorten geschiedt op verschillende manieren, afhankelijk van de beschikbare informatie.

Ten eerste kan men wegen naar diagnostische betekenis van de soorten door bijvoorbeeld kensoorten zwaarder laten wegen dan differentiërende soorten.

Ten tweede kan men de abundantie meer of minder laten meewegen. Dit is meestal van geringe betekenis: het feit dat de soort voorkomt is vaak veel belangrijker dan de mate waarin hij voorkomt.

Een derde type weging is naar de verwachting die kan worden gesteld aan het voorkomen van de soort in het syntaxon. Een soort die, hoewel kensoort, toch lang niet altijd wordt aangetroffen, kan hoger worden gewogen dan een kensoort die altijd aanwezig is. Voor deze weging kunnen de presentiewaarden van de soorten uit de samenvattende tabellen van de syntaxa worden gebruikt. Dit type weging is zeer belangrijk bij onevenwichtige indelingen, waarover later.

Het vierde type weging is naar het verschil tussen de abundantie en de verwachting die daaraan kan worden gesteld voor het syntaxon. Soorten met een hoge karakteristieke bedekking kunnen lager worden gewogen wanneer de abundantie in de opname gering is. Op deze manier kunnen met name dominantiegemeenschappen goed worden onderscheiden van soortenrijke gemeenschappen waarin dezelfde soorten een begeleidende rol spelen.

In de meeste syntaxonomische overzichten

worden naast kensoorten en differentiërende soorten ook andere aanduidingen gebruikt. Naast gewone (=selectieve) kensoorten worden ook exclusieve, preferente, lokale en zwakke kensoorten onderscheiden en naast gewoon ook zwak differentiërende soorten. Daarnaast wordt aan constante soorten een zekere diagnostische waarde toegekend en worden soorten vermeld die hun optimum of zwaartepunt hebben in het desbetreffende syntaxon.

Bij de preferente kensoorten zal een dubbele weging moeten plaatsvinden. De diagnostische waarde neemt toe bij een hogere bedekking. Ook bij kencombinaties moet een dubbele weging worden toegepast. De soorten uit zulke combinaties zijn slechts diagnostisch als ze (vrijwel) allemaal voorkomen.

Differentiërende soorten worden meestal beschreven ter onderscheiding van nevenschikte syntaxa, zodat ze geschikt zijn als criterium op het desbetreffende niveau. Soms worden soorten echter als differentiërend beschreven ten opzichte van verwante syntaxa uit een heel ander deel van de hiërarchische indeling. De waarde van zo'n aanduiding is betrekkelijk. Enerzijds kan zo'n diagnostische aanduiding de diagnose bevorderen als men inderdaad moet kiezen tussen deze syntaxa, maar anderzijds werden deze syntaxa ook al door kensoorten van hogere orde van elkaar onderscheiden, zodat het alleen zin heeft zulke differentiërende soorten te gebruiken als de relevante kensoorten ontbreken, of als van beide syntaxa een ongeveer even hoge diagnostische score is berekend.

Ook kan het voorkomen dat een differentiërende soort een bepaald syntaxon wel scheidt van het ene nevenschikt syntaxon, maar niet van het andere. Als voorbeeld kan een opname in een duinvallei van een duidelijk *Empetrium nigri* dienen. *Ammophila arenaria* is differentiërend voor

het *Polypodio-Empetretum* ten opzichte van de andere associaties, maar niet ten opzichte van het *Carici-Empetretum*. Aangezien de laatste associatie nauwelijks eigen diagnostische soorten heeft, behalve enkele lichenen die gemakkelijk foutief geïdentificeerd kunnen worden, kan het voorkomen dat het programma een opname met *Ammophila arenaria* louter om die reden diagnosticeert als een *Polypodio-Empetretum*. Diagnose vindt echter plaats op basis van meer dan één criterium, en in dit geval zal bij gebrek aan voldoende score van een der associaties, benoeming slechts op verbodsniveau het meest juist blijken.

Onevenwichtige indelingen

In ideale syntaxonisch overzichten hebben alle syntaxa evenveel kensoorten en differentiërende soorten, is de betekenis daarvan eenduidig en consequent, is het aantal associaties in alle klassen ongeveer gelijk en is de hiërarchische opbouw regelmatig. In de praktijk blijkt dat voor geen enkel syntaxonisch overzicht te gelden, zodat het programma de mogelijk moet hebben om op onevenwichtige indelingen in te spelen.

Sommige syntaxa hebben veel kensoorten maar er zijn ook associaties die helemaal geen eigen kensoorten hebben; sommige syntaxa, waaronder de meeste subassociaties worden slechts door (soms veel) differentiërende soorten van elkaar onderscheiden, maar bij *inops*-subassociaties ontbreken zelfs die.

Dit probleem wordt ondervangen door de verzadiging van een syntaxon zwaar te laten wegen. Onder verzadiging wordt daarbij verstaan het quotiënt van het aantal aangetroffen soorten van het syntaxon en het aantal te verwachten soorten van dat syntaxon. Daardoor kan de score van een syntaxon met maar één kensoort hoger uitvallen (als die soort aanwezig is) dan van

Diagnoses met totale scores:

Opname 1: Eleocharitetum multicaulis {06Ac03} (33)

Opname 2: Eleocharitetum multicaulis {06Ac03} (25)

Opname 3: Eleocharitetum multicaulis/Caricion nigrae {06Ac03/09Aa} (18/11)

Opname 4: Scirpetum fluitantis/Eleocharitetum multicaulis {06Ac02/06Ac03} (32/29)

Opname 5: Scirpetum fluitantis/Eleocharitetum multicaulis {06Ac02/06Ac03} (17/16)

Diagnostische score voor de belangrijkste syntaxa:

| | | 1 | 2 | 3 | 4 | 5 |
|--------|----------------------------|----|----|----|----|----|
| 06 | Littorelletea | 24 | 18 | 13 | 22 | 12 |
| 06A | Littorelletalia | 10 | 11 | 9 | 12 | 6 |
| 06Ac | Hydrocotylo-Baldellion | 20 | 18 | 14 | 22 | 11 |
| 06Ac02 | Scirpetum fluitantis | 6 | . | . | 12 | 6 |
| 06Ac03 | Eleocharitetum multicaulis | 10 | 11 | 9 | 9 | 5 |
| 09 | Parvocaricetea | 14 | 9 | 8 | 11 | 6 |
| 09A | Caricetalia nigrae | 5 | 3 | 3 | 5 | 3 |
| 09Aa | Caricion nigrae | 5 | 3 | 3 | 5 | 3 |
| 09Ba | Caricion davallianae | 1 | . | 1 | . | . |
| 11 | Oxycocco-Sphagnetea | 3 | . | . | . | 6 |
| 11A | Erico-Sphagnetalia | 3 | . | . | . | 6 |
| 11Aa | Ericion tetralicis | 3 | . | . | . | 6 |
| 11Aa01 | Lycopodio-Rhynchosporium | 4 | . | . | . | 7 |
| 16 | Molinio-Arrhenatheretea | . | 5 | 7 | . | . |
| 16A | Molinietalia | . | 10 | 11 | . | . |
| 16Aa | Junco-Molinion | . | 11 | 16 | . | . |
| 16Aa01 | Cirsio dissecti-Molinietum | . | 5 | 7 | . | . |

Syntaxonomische elemententabel:

| | | 1 | 2 | 3 | 4 | 5 |
|--------|----------------------------|----------|-----|-----|-----|---------|
| 06 | Littorelletea | | | | | |
| | Juncus bulbosus | zk.63.I | 2.2 | 1.2 | . | 2.2 2.2 |
| | Littorella uniflora | k.52.II | 3.3 | . | . | . |
| | Eleocharis multicaulis | T-k.45.I | +2 | 2.3 | 1.2 | +2 1.3 |
| | Echinodorus ranunculoides | k.27 | . | . | . | r . |
| 06A | Littorelletalia | | | | | |
| 06Ac | Hydrocotylo-Baldellion | | | | | |
| | Hypericum elodes | T-k.51.I | 1.2 | . | . | +2 1.2 |
| | Scirpus fluitans | T-k.41.I | . | . | . | 1.2 . |
| | Hydrocotyle vulgaris | * d.73.+ | 1.1 | +1 | 1.1 | 2.2 1.1 |
| | Ranunculus flammula | * d.55 | 2.2 | +1 | . | 1.1 . |
| | Agrostis canina | * d.26.+ | 2.2 | 1.2 | 1.2 | 3.2 +1 |
| 06Ac02 | Scirpetum fluitantis | | | | | |
| | Scirpus fluitans | k.90.II | . | . | . | 1.2 . |
| | Hypericum elodes | k.78.I | 1.2 | . | . | +2 1.2 |
| 06Ac03 | Eleocharitetum multicaulis | | | | | |
| | Eleocharis multicaulis | k.98.I | +2 | 2.3 | 1.2 | +2 1.3 |
| | Deschampsia setacea | k.71 | 2.2 | 1.2 | 1.2 | +2 . |
| 09 | Parvocaricetea | | | | | |
| | Hydrocotyle vulgaris | k.58.+ | 1.1 | +1 | 1.1 | 2.2 1.1 |
| | Ranunculus flammula | k.31 | 2.2 | +1 | . | 1.1 . |
| | Calliergonella cuspidata | lk.45.II | +1 | +1 | +1 | . . |
| | Potentilla palustris | T-k.42.I | +1 | . | . | +1 . |

| | | | | | | |
|-----------------------------------|----------|-----|-----|-----|-----|-----|
| Agrostis canina | T-k.35.I | 2.2 | 1.2 | 1.2 | 3.2 | +1 |
| Carex nigra | T-k.30.+ | +1 | +1 | 1.1 | . | +1 |
| Galium palustre | d.39 | +1 | . | 1.1 | . | . |
| Lythrum salicaria | * d.34 | 2.2 | +1 | . | 1.1 | . |
| 09A Caricetalia nigrae | | | | | | |
| 09Aa Caricion nigrae | | | | | | |
| Potentilla palustris | k.53.I | +1 | . | . | +1 | . |
| Agrostis canina | k.52.I | 2.2 | 1.2 | 1.2 | 3.2 | +1 |
| Carex nigra | k.42.I | +1 | +1 | 1.1 | . | +1 |
| 09Ba Caricion davallianae | | | | | | |
| Salix repens | d.64.I | 1.2 | . | 2.2 | . | . |
| 11 Oxycocco-Sphagnetea | | | | | | |
| 11A Erico-Sphagnetalia | | | | | | |
| 11Aa Ericion tetralicis | | | | | | |
| Carex panicea | d.31 | . | +2 | 2.2 | . | . |
| Gentiana pneumonanthe | * d.20 | . | +1 | 1.2 | . | . |
| 11Aa01 Lycopodio-Rhynchosporium | | | | | | |
| Drosera intermedia | k.82.+ | r | . | . | . | +1 |
| Lycopodium inundatum | k.53.+ | . | . | . | . | r |
| Molinia caerulea | * c.90.I | +2 | +2 | 2.2 | . | 2.2 |
| 16 Molinio-Arrhenatheretea | | | | | | |
| 16A Molinieta | | | | | | |
| Galium palustre | d.41 | +1 | . | 1.1 | . | . |
| Calliergonella cuspidata | * d.42.I | +1 | +1 | +1 | . | . |
| Carex nigra | * d.32.+ | +1 | +1 | 1.1 | . | +1 |
| Hydrocotyle vulgaris | * d.34.+ | 1.1 | +1 | 1.1 | 2.2 | 1.1 |
| Lythrum salicaria | * d.37 | 2.2 | +1 | . | 1.1 | . |
| Lysimachia vulgaris | d.34 | . | +1 | . | +1 | . |
| 16Aa Junco-Molinion | | | | | | |
| Molinia caerulea | d.93.I | +2 | +2 | 2.2 | . | 2.2 |
| Potentilla erecta | d.80 | . | . | +1 | . | . |
| Gentiana pneumonanthe | * d.31 | . | +1 | 1.2 | . | . |
| Agrostis canina | * d.60.+ | 2.2 | 1.2 | 1.2 | 3.2 | +1 |
| 16Aa01 Cirsio dissecti-Moliniatum | | | | | | |
| Cirsium dissectum | k.81.+ | . | +1 | 2.2 | . | . |
| Carex panicea | zk.89.+ | . | +2 | 2.2 | . | . |
| Overige soorten | | | | | | |
| Carex oederi | | 2.2 | 2.3 | . | 1.2 | +2 |
| Drepanocladus sendtneri | | 2 | . | . | . | . |
| Leontodon autumnalis | | . | +1 | . | . | +1 |
| Mentha arvensis | | 1.2 | +1 | . | . | . |
| Potentilla anserina | | +1 | +1 | . | . | . |
| Salix spec. | | . | +1 | . | . | +1 |
| Sphagnum subsecundum | | . | +1 | 1.2 | . | 4.4 |

Tabel I. Analyse van de opnamen van Hofstra (1982), genoemd in Schaminée et al. (1995b): p. 126. Verklaring voor de codering in de achter de soortnamen: T=transgrediërend; *=ook elders in de tabel; (l/z)k=(lokale/zwakke) kensoort; d=differentiërende soort; c=constante soort; getal=presentie; Romeinse cijfer= karakteristieke bedekking (ontleend aan Schaminée et al. 1995c).

een syntaxon waarvan maar 5 van de 9 kensoorten aanwezig zijn. Het aantal te verwachten kensoorten wordt, zo mogelijk, echter bepaald door de som van de presentiewaarden in de syntaxa van deze kensoorten. Als van die genoemde 9 kensoorten er 7 zeer zeldzaam zijn, met een presentie van gemiddeld 15%, dan is de verwachting slechts ongeveer 3, en scoort dit syntaxon met een verzadiging van 5/3 toch hoger dan het eerste.

Een vergelijkbaar diagnoseprobleem doet zich ook voor bij de keuze tussen een klasse met veel hiertoe behorende syntaxa en een met weinig. Bij een klasse met veel associaties kan het gebeuren dat het totaal aantal kensoorten dat bijdraagt aan de score voor de klasse weliswaar groot is, maar dat deze verspreid zijn over alle associaties van de klasse. Geen der associaties is dan sterk vertegenwoordigd en diagnose van een associatie uit een heel andere klasse kan wellicht correcter zijn. De oplossing voor dit probleem wordt gevonden bij de doorwerking van de diagnostische score van elk syntaxon naar bovenliggende syntaxa. Bij de score van een verbond wordt niet de totale score van de kensoorten van alle daartoe behorende associaties bijgeteld, maar slechts die van de hoogst-scorende associatie. Op die manier maakt het niet uit hoeveel associaties een verbond heeft voor de bijtelling en wordt de hoogst scorende associatie toch wel op de juiste waarde geschat. Hetzelfde geldt uiteraard ook voor hogere niveaus.

Een speciaal, hoewel veel voorkomend geval is dat waarbij syntaxa slechts één onderliggend syntaxon hebben. Veelal worden in de beschrijving helemaal geen kensoorten vermeld en in de samenvattende tabel wordt slechts één kolom gegeven voor beide niveaus. Wanneer kensoorten op beide niveaus zouden worden gerekend, zou dat een onevenredige score op klasse-ni-

veau geven. In zo'n geval wordt altijd op het lagere niveau geïdentificeerd. Als bijvoorbeeld de klasse *Trifolio-Geranietea sanguinei* het hoogste scoort, dan is ook meteen het verbond *Trifolion medii* geïdentificeerd. Iets dergelijks geldt ook voor gemeenschappen zonder diagnostische soorten (*inops*, *typicum*); deze worden als zodanig geïdentificeerd bij afwezigheid van diagnostische soorten van alle neven-geschiedte syntaxa.

Onvolkomen diagnose: romp- en derivaatgemeenschappen

Diagnostische soorten van lagere syntaxonomische niveaus zijn in de regel het meest kritisch ten aanzien van hun milieueisen en zullen onder allerlei storingsomstandigheden het eerst ontbreken. Dat betekent dat dan de diagnose voor een associatie lastig is te bereiken en men de conclusie moet trekken dat men met een verarmde gemeenschap heeft te maken, die afhankelijk van de dominante soort of soorten romp- of derivaatgemeenschap kan worden genoemd. Bij elke beoordeling welk syntaxon *het hoogst* scoort op een bepaald niveau moet ook worden geoordeeld of de score wel *hoog genoeg* is voor identificatie. Om te bepalen worden (instelbare) criteria gehanteerd.

I. Kopecký et al. (1995) stellen als criterium dat er voldoende diagnostische soorten moeten zijn ten opzichte van de hogere niveaus volgens een bepaalde regel. Omdat in de scores op de hogere niveaus ook de kensoorten op lagere niveaus zijn opgenomen, neemt de score naar beneden in de regel af. Als die hogere niveaus geen eigen score hebben zal de score van de lagere niveaus door differentiërende soorten echter ook hoger kunnen zijn. De vraag is hier: welke scoreverhouding is minimaal nodig voor verdere diagnose. Kopecký et al. hou-

den hier voor het syntaxonomisch overzicht waarmee zij werkten de factor 1,6 aan.

Dit criterium blijkt echter heel gevoelig te zijn voor een onevenwichtige opbouw van het syntaxonomische overzicht wat betreft het aantal diagnostische soorten per syntaxon. Syntaxa met weinig of geen diagnostische soorten op een laag niveau en veel diagnostische soorten op een hoog niveau worden dan niet gemakkelijk geïdentificeerd, terwijl het toch een goed ontwikkelde gemeenschap betreft. Dit treedt met name op bij verbonden met een redelijk groot aantal diagnostische soorten, waartoe associaties behoren met op hun beurt heel weinig diagnostische soorten. Het quotiënt van de score van het verbond en van de associatie is dan veelal veel te hoog.

2. Als van het te beoordelen niveau de verzadiging hoog genoeg is, dat wil zeggen dat een substantieel deel van te verwachten diagnostische soorten aanwezig is, dan is de identificatie duidelijk. Dit kan onafhankelijk van het niveau worden getoetst, zodat bovenstaand probleem in principe kan worden omzeild. De vraag is dan natuurlijk: welke minimale verzadiging moet worden aangehouden?

Dit criterium is echter op een ander punt gevoelig voor het aantal diagnostische soorten en hun zeldzaamheid. Van sommige syntaxa zijn veel diagnostische soorten beschreven, zodat verzadiging niet hoog hoeft te zijn om te weten dat het om desbetreffende syntaxa gaat. In allerlei testen bleek dit verzadigingscriterium op zichzelf heel slecht te werken.

In combinatie met het eerste criterium blijkt het wel goede mogelijkheden te hebben, omdat de bezwaren bij complementaire situaties optreden.

3. Als diagnostische soorten van hogere niveaus (vrijwel) ontbreken, maar er zijn voldoende associatie-kensoorten, mag dit de identificatie van de associatie niet beletten, al zal men in zo'n geval soms van

associatiefragment willen spreken. In de beoordeling volgens het eerste criterium wordt dan niet alleen gekeken naar de verhouding tussen het reeds geïdentificeerde niveau en het eerstvolgende daaronder, maar ook tussen het eerste en nog lagere. Bij vergelijking over meer dan één niveau wordt echter alleen de bijdrage van de kensoorten aan de score van het lagere niveau bij de vergelijking betrokken.

4. Als het syntaxon op het desbetreffende niveau het enige syntaxon is, of het is een *inops*- of *typicum*-gemeenschap zonder diagnostische soorten terwijl diagnostische soorten van de nevenschikte syntaxa in de opname ontbreken, dan kan identificatie ook plaatsvinden.

5. Stopt de afdaling in de hiërarchie op een hoger niveau dan associatie, dan betreft de onderzochte opname een romp- of derivaatgemeenschap.

Overgangen

Heel vaak zal het gebeuren dat benoeming wel kan plaatsvinden, wellicht zelfs tot op associatieniveau, maar dat ook aspecten van een of twee andere syntaxa in zodanig mate aanwezig zijn dat benoeming tot die syntaxa ook met recht zou kunnen plaatsvinden. Dit betreft dan een overgang (ruimtelijk of temporeel, gradiënt, mozaïek, schering en inslag, verstoring, ontwikkelingsfase, degeneratiefase, enz.). Er zal dan altijd een eerste diagnose en ten minste één alternatief moeten worden gegeven, om de onderzoeker de kans geven deze tegen elkaar af te wegen. Ook moet de vegetatie als overgang kunnen worden geïdentificeerd. Om het aantal alternatieven te beperken moet de verhouding tussen de score van de eerste keuze en de minimale score voor de tweede en volgende keuzes worden ingesteld.

Afhankelijk van de doelstelling van het onderzoek kan men verschillend omgaan

met overgangen:

- Bij de identificatie van individuele opnamen zal men de soorten die daarbij een rol spelen willen vergelijken en de complete beschrijvingen in de publicatie erover moeten raadplegen, teneinde de meest waarschijnlijke syntaxonomische positie te kunnen vaststellen.
- Bij vergelijking van opnamen in het zelfde gebied (ruimtelijke variatie) of in de tijd (permante quadraten) kan men de diagnostische scores voor elk syntaxon met elkaar vergelijken en conclusies trekken over toename en afname van de bijdrage van de afzonderlijke syntaxa.

Similariteit als verificatie

Bij overgangen en bij onduidelijke identificatie kunnen begeleidende soorten een belangrijke rol spelen. De diagnostische soorten van één van de alternatieven kunnen wellicht wel als begeleidend bij de ander zijn beschreven, maar niet of in veel mindere mate andersom. Om deze vergelijking te maken wordt een similariteit-berekening uitgevoerd volgens Sørensen. De Jaccard-index is hier minder geschikt, omdat deze in de berekening alle betrokken soorten volledig meetelt. Omdat de lijst met soorten van de syntaxa vrijwel altijd (veel) langer is dan die van de opname, zal dat doorgaans tot onevenredig lage scores leiden. De Sørensen-index wordt zo mogelijk kwantitatief berekend op analoge wijze als de Ball-index (zie voor een bespreking van deze indices Hennekens et al. 1995, pg. 154).

Ook als samenvattende tabellen ontbreken kan de berekening worden uitgevoerd, maar dan met alle diagnostische soorten van het betreffende en alle hogere niveaus.

Het kan gebeuren dat de scores voor de diagnose van twee of meer syntaxa elkaar niet zo ver ontlopen, maar dat de similariteits-indices voor deze syntaxa wel sterk

verschillen. Dit treedt onder andere op bij de vergelijking van een soortenarm syntaxon met een soortenrijk syntaxon waarbij de kensoorten van het eerste ook als begeleider in het tweede optreden. Een voorbeeld hiervan betreft *Phragmites australis*, preferente kensoort van het *Typho-Phragmitetum*, die als successierelict nog lang, zelfs aspectbepalend kan standhouden in de *Filipenduletea*, het *Calthion palustris* en de *Parvocaricetea*, die alle veel soortenrijker zijn. Diagnose van een rietlandopname zou in eerste instantie al deze alternatieven kunnen opleveren. Door de similariteit te berekenen komt het verschil aan het licht: de vegetatie kan tot de *Typho-Phragmitetum* worden gerekend als ze soortenarm is, maar moet tot een der laatste syntaxa worden gerekend als ze soortenrijk is (Schaminée et al. 1995c).

Een bijzondere mogelijkheid doet zich voor bij de verificatie van een diagnose waarbij een romp- of derivaatgemeenschap is gevonden en er in het syntaxonomisch overzicht ook algemeen voorkomende romp- en derivaatgemeenschappen worden beschreven, zoals bij Schaminée et al. (1995c, 1996, in prep). Er kan dan worden bezien of er een romp- of derivaatgemeenschap is beschreven onder dezelfde klasse met dezelfde dominante soort of soorten, waarmee de similariteit groter is dan van het geïdentificeerde syntaxon zelf. Wanneer een opname bijvoorbeeld wordt geïdentificeerd als RG *Holcus lanatus*-[*Molinietalia*], dan kan de vergelijking tussen de similariteits-indices van de opname met RG *Holcus lanatus*-*Lychnis flos-cuculi*-[*Molinietalia*], met RG *Holcus lanatus*-*Lolium perenne*-[*Molinio-Arrhenatheretea*] en met de *Molinietalia* een idee geven of men met een bekende dan wel nieuwe rompgemeenschap te maken heeft.

Aan het gebruik van de samenvattende ta-

bellen uit een syntaxonomisch overzicht kleeft nog een bezwaar. Er bestaat nog wel eens discrepantie tussen de opsomming van diagnostische soorten in de tekst en de cijfers in samenvattende tabellen die deze beschrijving zouden moeten ondersteunen. Hoewel bij het onderscheiden van diagnostische soorten de tabellen het voornaamste hulpmiddel vormen, blijft altijd nog een beoordeling aan de hand van vegetatiekundige veldervaring en literatuurkennis nodig. Daaruit komt voort dat diagnostische soorten soms ten onrechte als zodanig lijken te zijn opgevoerd. Een bekend voorbeeld is *Caltha palustris* in het *Calthion*. De soort is volgens sommigen nauwelijks diagnostisch voor dat verbond op grond van de presentiecijfers, maar voldoet in de praktijk toch uitstekend als diagnostische soort. Moeilijk te kwantificeren informatie of van vergelijkbaar buitenlands materiaal afgeleide kennis heeft vaak een rol gespeeld bij het benoemen van de diagnostische soorten. Dit komt bij talrijke klassen in Schaminée et al. (1995c, 1996) voor. In overzichten zonder samenvattende tabellen zoals dat van Westhoff & Den Held (1969) is dit niet zichtbaar, maar ook hier heeft interpretatie in groter verband ongetwijfeld wel eens noodgedwongen tot interpretatieverschillen met de oorspronkelijk (veelal door anderen) gepubliceerde tabellen geleid.

Tenslotte

Identificatie met behulp van de computer is waarschijnlijk nooit volkomen sluitend te krijgen. In sommige gevallen is aanvullende informatie over fenologische toestand van de soorten nodig, bijvoorbeeld om onderscheid te kunnen maken tussen struweelgemeenschappen en bossen met sterk overeenkomstige soortensamenstelling, zoals respectievelijk de *Rhamno-Prunetea* met sommige vormen van de *Quercu-Fage-tea* (mond. med. R. Haveman).

Ook is het goed denkbaar dat identificatie plaats heeft op basis van een zeer gering aantal soorten, omdat ook van de eventuele alternatieven maar weinig soorten aanwezig zijn.

Daarentegen komt het ook vaak voor dat identificatie juist niet lukt door determinatieproblemen. Als lichenen, *Callitriche*-soorten, *Rubus*-soorten, *Taraxacum*-soorten of bepaalde subspecies niet juist of helemaal niet op naam zijn gebracht wordt een diagnostische soort niet als zodanig herkend.

Dit herkenningsprobleem wordt in principe nog groter als de soortcodering van de opnamen niet overeenkomt met die van het syntaxonomisch overzicht, maar dat kan technisch wel worden opgevangen met de huidige programmeertechnieken.

Controle achteraf van alle soorten in hun syntaxonomisch perspectief is daarom een vereiste. Het is ondenkbaar dat een onderzoeker met een 'druk op de knop' een diagnose stelt zonder zich af te vragen hoe die diagnose tot stand is gekomen.

SynDiaT levert daartoe aan het eind van de berekeningen tabellen met diagnostische scores, verzadigingspercentages, similariteits-indices, diagnostische soorten en syntaxonomische elementen. In de diagnostische soortentabel worden alle gevonden syntaxa opgesomd met de erbij behorende diagnostische soorten en is vooral bedoeld voor het beoordelen van de wijze waarop de diagnose tot stand kwam. De syntaxonomische elemententabel is een vereenvoudiging hiervan, waarbij alle syntaxa zijn weggelaten waarvoor in geen der onderzochte opnamen een voldoende score is gevonden voor identificatie. Deze tabel geeft alleen de meest relevante syntaxonomische positie(s) van de soorten in de opnamen, is daardoor gemakkelijker leesbaar en kan tevens als basis dienen voor een te publiceren tabel.

Voorbeeld

Ter afsluiting de uitwerking van de diagnose van het voorbeeld dat ook door Schaminée et al. (1995b) werd gebruikt. Duide-lijk komt uit de diagnose (tabel I) naar voren dat het hier het *Hydrocotylo-Baldellion* be-treft. Ook wordt duidelijk dat de eerste drie opnamen als *Eleocharitetum multicaulis* en de laatste twee als een overgang tussen deze en het *Scirpetum fluitantis* kunnen worden beschouwd. In het tweede deel van de tabel is voor alle opnamen de hoogste klasse-sco-re bij de *Littorelletea* te vinden. Alternatie-ven worden op elk niveau gegeven als ze meer dan 60% van de hoogste score halen. In het derde deel is te zien door welke soorten deze diagnose tot stand komt. Op-vallend is de beoordeling van de rest van de soorten in van de tabel. In de oorspronkelijk-tabel van Hofstra (1982) is de aandacht gericht op *Junco-Molinion/Nardo-Galion saxatilis*, *Lycopodio-Rhynchosporium* en *Caricion davallianae*. In tabel I is echter duidelijk te zien dat deze diagnose helemaal niet zo sterk is en merendeels is gebaseerd op wat volgens Schaminée et al. (1995c) diffe-rentiërende soorten zijn, en dat het *Caricion nigrae* toch met minstens zoveel recht genoemd had kunnen worden. In de derde opname wordt het laatste zelfs g-eïdentificeerd als overgang (indringing), overigens vooral omdat de *Littorelletea* hier relatief laag scoren en daardoor de *Parvo-caricetea*-score uitkomt boven 60% van de *Littorelletea*-score.

Identification of vegetation types by using computer programs

In contrast to classification and ordination hardly any successful computer program has been developed for identification of vegetation types so far. Most attempts are based on comparison of the species compo-sition of relevés with that of all types in a

given classification. In many cases results are disappointing if no complete listing of the classification is given. In most relevés the majority of species have broad ecologi-cal amplitudes which usually makes the relevés almost equally resembling to many types.

On the other hand, analyses with diag-nostic species require a lot of experience and knowledge of the classification invol-ved. Most of the diagnostic species also occur in other types than those they charac-terize. Therefore identification is a complex process and rules are difficult to be formal-ized into a computer program.

Kopecký et al. (1995) published an al-gorithm for deductive analyses with diag-nostic species, which has served as a basis for the computer program SynDiaT. Basic-ally the program calculates the scores of relevés for every syntaxon in a given clas-sification by counting diagnostic species in the relevés. All scores are weighed for a variety of factors. Character species are considered to be diagnostic for all higher levels in the hierarchy as well.

Identification takes place by searching top-down through the hierarchy for the hig-hest scores. Adaptations have been develo-ped for unbalanced systems and rules are given for identification of basal and deriva-tive communities, both in which saturation indices are included. Identification of se-cond and other alternatives are given within thresholds.

Sørensen indices of similarity for the relevés with every syntaxon identified is used for verification of the results and re-consideration of the alternatives.

Computer aided identification of vege-tation types is not believed to become per-fect anyway. Certain problems most likely cannot be solved at all. SynDiaT therefore makes a complete report with all relevant data and all diagnostic species per syntaxon identified to be used for a final reconsidera-

tion 'by hand'.

Gerefereerde literatuur

- Hennekens, S.M., E. van der Maarel & A.H.F. Stortelder (1995). Numerieke methoden. In: Schaminée, J.H.J., A.H.F. Stortelder & V. Westhoff (red.); *De Vegetatie van Nederland I Inleiding in de plantensociologie - grondslagen, methoden en toepassingen*. Opulus Press, Uppsala en Leiden, pp. 151-162.
- Hill, M.O. (1993). TABLEFIT - For identification of vegetation types. Report Institute of Terrestrial Ecology, Monks Wood.
- Hofstra, J. (1982). Over enige Littorelleteagemeenschappen, in het bijzonder in Twente. *Gorteria* 11: 59-72.
- Kopecký, K. & S. Hejný (1974). A new approach to the classification of anthropogenic plant communities. *Vegetatio* 29: 17-20.
- Kopecký, K., J. Dostálek & T. Frantík (1995). The use of the deductive method of syntaxonomical classification in the system of vegetational units of the Braun-Blanquet approach. *Vegetatio* 117: 95-112.
- Pot, R. (1997). SYNDIAT, SYNtaxonomical DIAgnostics Tool, a computer program based on the deductive method of community identification. *Acta Botanica Neerlandica* 46: 230.
- Schaminée, J.H.J., A.H.F. Stortelder & V. Westhoff (1995a). *De Vegetatie van Nederland 1. Inleiding in de plantensociologie - grondslagen, methoden en toepassingen*. Opulus Press, Uppsala en Leiden. 296 pp.
- Schaminée, J.H.J., A.H.F. Stortelder & J.J. Barkman (1995b). Syntaxonomische identificatie. In: Schaminée, J.H.J., A.H.F. Stortelder & V. Westhoff (red.); *De Vegetatie van Nederland 1*. Opulus Press, Uppsala en Leiden, pp. 115-128.
- Schaminée, J.H.J., E.J. Weeda & V. Westhoff (1995c). *De Vegetatie van Nederland 2. Wateren, moerassen en droge heiden*. Opulus Press, Uppsala en Leiden. 358 pp.
- Schaminée, J.H.J., A.H.F. Stortelder & E.J. Weeda (1996). *De Vegetatie van Nederland 3. Graslanden, zomen en droge heiden*. Opulus Press, Uppsala en Leiden. 356 pp.
- Westhoff, V. & A.J. Den Held (1969). *Plantengemeenschappen in Nederland*. Thieme, Zutphen. 324 pp.