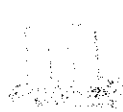


ppr

11-10-12



Gez. BROTHIER
RPO sector Fruit
Postbus 200
2070 AH Zetten
0488 478700



CURSUS ONDERZOEKTECHNIEK

Syllabus: Classificatie, toegelicht aan voorbeelden uit de
Veldbodemkunde.

Dr.ir.J.Schelling, Stichting voor Bodemkartering.

1. INLEIDING

1.1 Algemene ontwikkeling van een wetenschap

Iedere wetenschap doorloopt een ontwikkeling, die globaal gezien in een aantal fasen is te verdelen. Dit zijn achtereenvolgens:

- beschrijving
- classificatie
- inventarisatie
- verklarend onderzoek en theorievorming.

In werkelijkheid verloopt de ontwikkeling niet precies volgens deze indeling, maar vindt de ontwikkeling van meerdere fasen gelijktijdig plaats, waarbij bepaalde onderdelen binnen een wetenschap al veel verder voortgeschreden kunnen zijn dan andere.

De mate van precisie neemt meestal toe bij voortschrijdende ontwikkelingen, waarbij men aanvankelijk bij beschrijving en classificatie kwalitatief te werk gaat, om vervolgens meer kwantitatieve methoden te gaan hanteren.

1.2 Veldbodembkunde

Onder veldbodembkunde verstaan we dat gedeelte van de bodembkunde dat speciaal gericht is op de kennis van de bodem zoals die in het veld voorkomt.

De bodem is het bovenste gedeelte van de aardkorst voor zover deze door bodemvormende processen is veranderd (zoals b.v. een humeuze bovengrond) of voor zover deze door planten kan worden beworteld.

Uit deze definitie blijkt reeds, dat het begrip bodem voor de meer theoretische, op het ontstaan gerichte, bodembkunde wat anders is dan voor de op de toepassingen voor de planteteelt gerichte bodembkunde.

De bodem als object van de veldbodembkunde heeft een aantal kenmerken die de aanpak van beschrijving, classificatie en verklarend onderzoek in sterke mate bepalen.

Dat zijn de volgende:

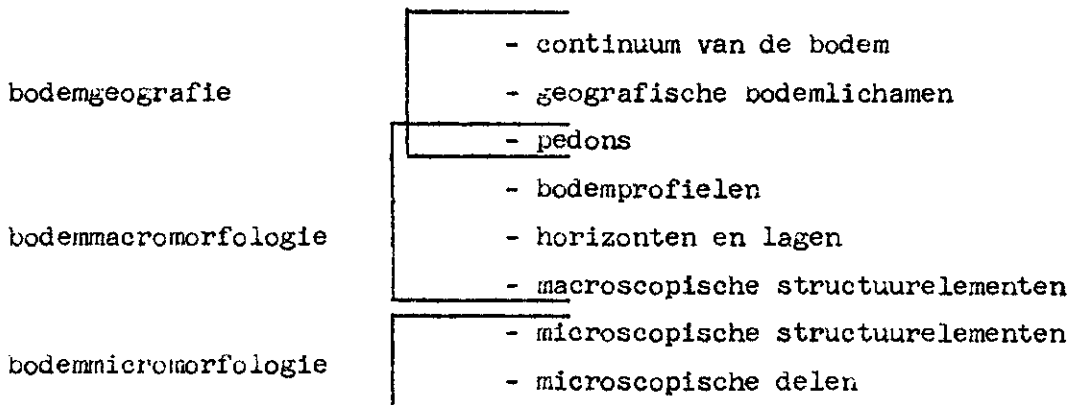
- het object ligt in het veld;
- het is uitgebreid, dat brengt o.a. mee dat er allerlei geografische aspecten aan zitten;
- het is moeilijk waar te nemen, direct waarneembaar is alleen het oppervlak; willen we de diepte in dan moeten we graven of boren, en zien toch nog maar een zeer klein gedeelte;
- de bodem is een continuum, d.w.z. dat er geen natuurlijke individuen zijn te onderscheiden, zoals b.v. een plant of een dier in de biologie; m.a.w. de bodem is één doorlopende schil, die van plaats tot plaats in eigenschappen verschilt, maar waar we niet gemakkelijk één bodem, één driedimensionaal stuk, in kunnen onderscheiden op een eenduidige wijze;
- het is maar in beperkte mate geschikt voor proeven met ingreep. Veel onderzoek zal daarom als vergelijkend onderzoek moeten worden aangepakt.

2. BESCHRIJVING

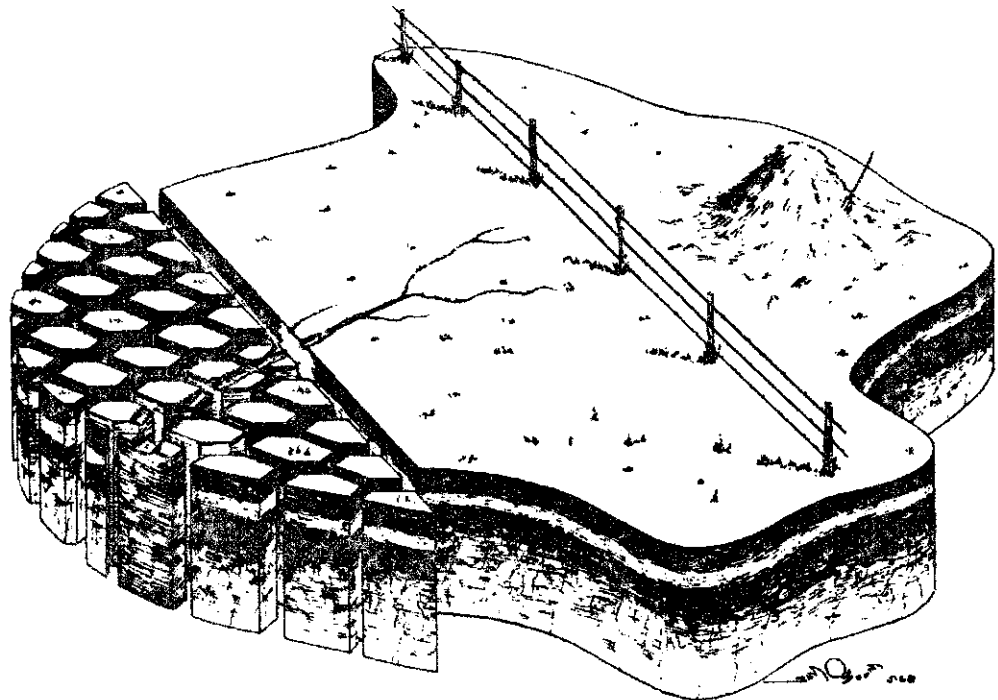
Wat beschrijven we, hoe beschrijven we dat en waarom beschrijven we het.

2.1 Wat beschrijven we in de veldbodemkunde

Een complex object zullen we in onderdelen moeten opdelen om het te kunnen beschrijven. Hiervoor kunnen we b.v. het volgende schema (zie fig. 1) gebruiken:



Op zich zelf vormt dit schema een classificatiesysteem.



Geografisch bodemlichaam

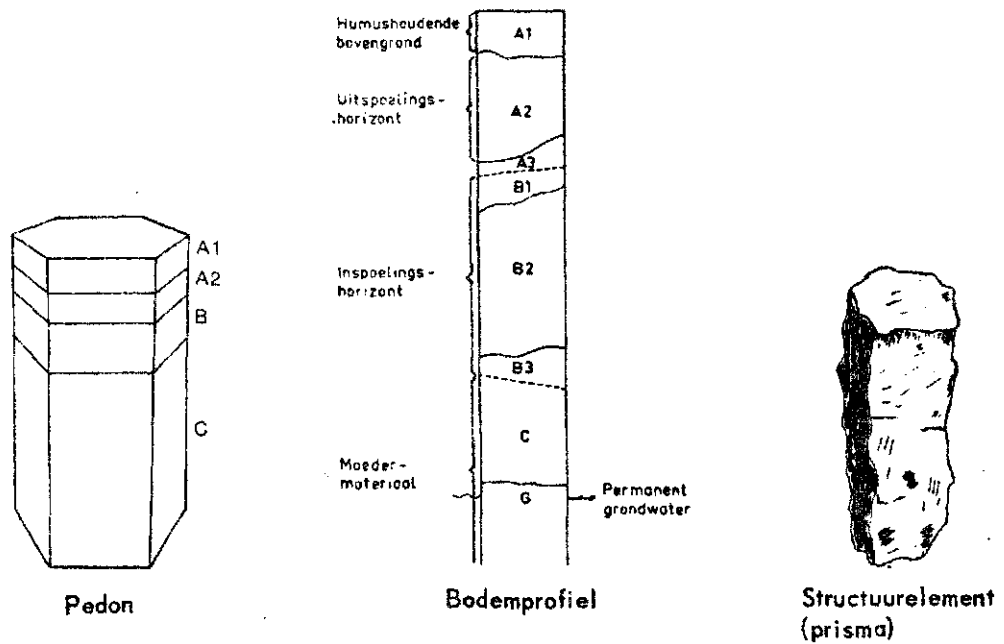


Fig. 1 Schematische voorstelling van enige onderdelen die worden gebruikt bij de beschrijving in de veldbodemkunde.

Dit schema heeft een bepaald karakter, het is morfologisch (morfologie is vormleer, in de biologieleer van de vorm en bouw der organismen). De bovenste categorieën (afdelingen) hebben grote afmetingen, ze behoren tot de bodemgeo-

grafie. De middelste hebben afmetingen die met ongewapende zintuigen zijn waar te nemen en de onderste zijn alleen microscopisch waarneembaar.

Er oestaat hier een duidelijke relatie tussen de aard van het systeem en van de onderdelen en de wijze van waarnemen.

Bestuderen we de bodem b.v. vanuit de natuurkunde (fysica) dan is de bodem samengesteld uit gas, vloeistof en vaste delen. Wanneer we hiervan de vloeistof beschouwen, die we eenvoudigweg als water aanduiden, dan is er afhankelijk van de methode van meten een indeling te maken.

Zetten we een buis van 5 cm doorsnede met geperforeerde wanden in de grond, dan kunnen we in natte gronden hierin een grondwaterstand meten. Verfolgen we deze grondwaterstand gedurende een jaar, dan is deze in een grafiek af te beelden als een z.g. tijdstijghoogtecurve. De grondwaterstand is dus een veranderlijke eigenschap. Herhalen we deze metingen gedurende een reeks van jaren, dan zijn deze curven verschillend, afhankelijk van externe omstandigheden (zoals het weer). Wat de eigenschap is die wordt ingedeeld blijkt nauw met de methode van meten samen te hangen.

2.2 Hoe beschrijven we

Aan verschillende onderdelen kunnen we eigenschappen gaan bepalen.

Hierbij moet worden vastgesteld:

- de bepalingsmethode
- de meeteenheden
- statistische eisen
- de classificatie van de eigenschap.

Dit wordt toegelicht aan de hand van een voorbeeld, de korrelgroottesamenstelling.

Textuurindeling

De classificatie van grond naar de korrelgroottesamenstelling wordt in het vakjargon textuurindeling genoemd. Deze geschiedt naar gewichtsprocenten "minerale delen", d.w.z. delen < 2 mm, bepaald volgens een voorgeschreven methode.

Door middel van bepaalde slib- en zeefmethoden wordt een grondmonster gescheiden in een aantal delen, de z.g. fracties, met voor iedere fractie exact aangegeven grenzen. De gebruikte meeteenheid is de μ , 1/1000 mm.

Voor een indeling kiezen we nu bepaalde grenzen uit die van belang zijn:

- materiaal $< 2 \mu$, lutumfractie, in het gewone spraakgebruik klei;
- materiaal $2-50 \mu$, de siltfractie;
- materiaal $50-2000 \mu$, de zandfractie.

De lutumfractie is van belang omdat we hierin materiaal vinden met een groot oppervlak, waar tal van chemische en fysische eigenschappen van de grond door worden bepaald.

De silt- en lutumfractie, te zamen als leemfractie aangeduid, bestaan uit materiaal dat gemakkelijk hoog door de lucht kan worden verplaatst. De löss bestaat voor een belangrijk deel uit deze fractie.

Het zand vormt het stabiele skelet van de grond en is materiaal dat bij verplaatsing door de wind laag over de grond wordt voortgewogen.

Voor een indeling van door de wind afgezet (z.g. eolisch) materiaal is in fig. 2 met behulp van bovengenoemde fracties een indeling gemaakt.

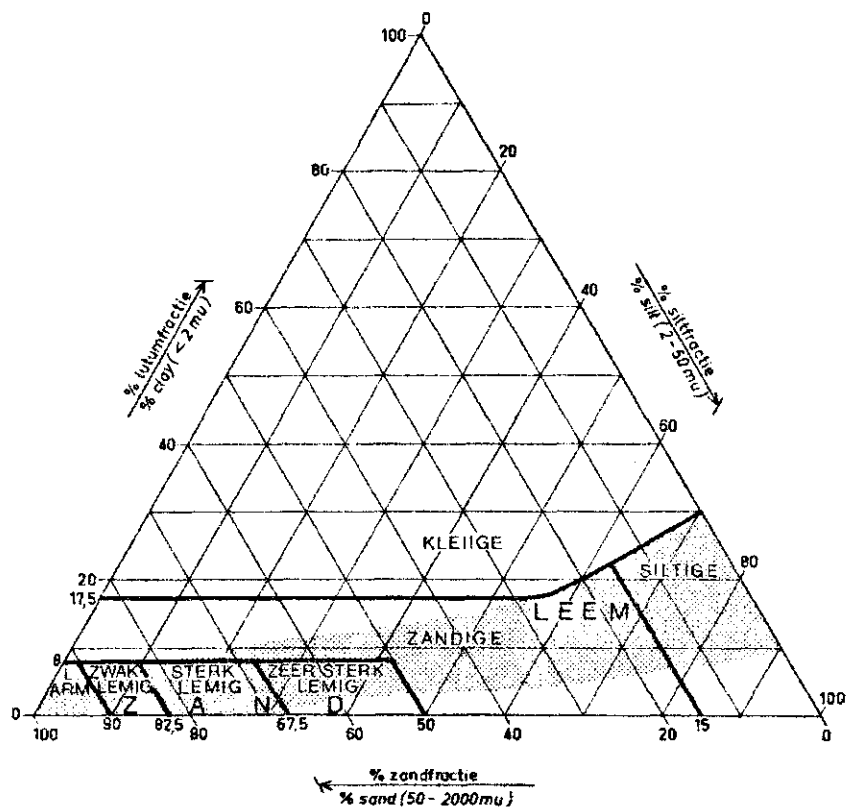


Fig. 2 Indeling en benaming van de textuur van eolische afzettingen met behulp van drie fracties.

De grijze zone geeft aan waar het merendeel van de Nederlandse windafzettingen voorkomt. De indeling is in hoofdzaak gebaseerd op het zandgehalte of het leemgehalte. Slechts één grens in het grijze gebied is op het kleigehalte (8 %) gebaseerd.

Uit dit voorbeeld blijkt dat het hoe van de indeling nauw samenhangt met het waarom.

De beschrijving van complexe onderdelen

Wat de morfologie betreft dient men hier meestal bij te onderscheiden:

- | | | |
|--------------|---|--|
| - vorm |) | |
| - grootte |) | van de samenstellende elementen |
| - aard |) | en hun onderlinge <u>rangschikking</u> |
| - omgrenzing |) | |

Dit wordt toegelicht aan de hand van het "structuurprofiel".

Het structuurprofiel

Bij de indeling van de macrostructuren hanteert men o.a. de volgende kenmerken:

1. vorm van de structuurelementen
2. afmetingen van de structuurelementen
3. structuurgraad, ruwweg aan te duiden als de mate van ontwikkeling.

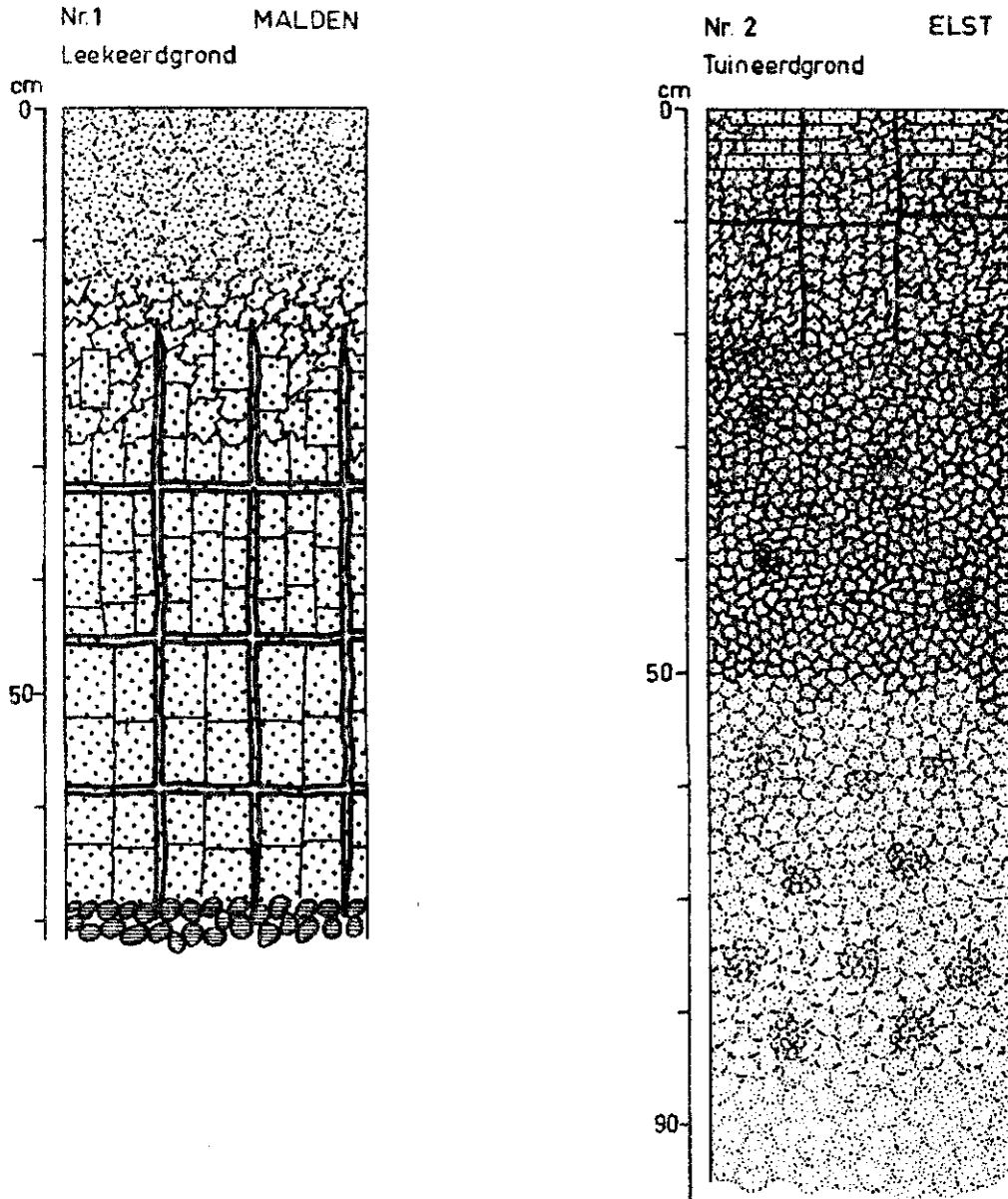


Fig. 3 Een tweetal structuurprofielen, voornamelijk ontstaan resp. onder invloed van zwel en krimp (nr. 1) en onder invloed van biologische activiteit (nr. 2).

1. De vorm

Hierin zijn drie groepen te onderscheiden;

A - vormen waarbij de drie assen nagenoeg gelijk zijn, o.a. blokkige elementen (fig. 3, nr. 1, laag 0-15 cm en nr. 2 in het gehele profiel)

B - vormen waarbij de verticale as langer is dan de horizontale assen (fig. 3, nr. 1, laag 15-70 cm), de prisma's

C - vormen waarbij de horizontale assen langer zijn dan de verticale, platige elementen.

Binnen deze groepen worden vervolgens een aantal soorten vormen onderscheiden. Op een typisch aspect van deze indeling wordt hier de aandacht gevestigd, nl. het onderscheid tussen samengestelde en enkelvoudige elementen.

In fig. 3, nr. 1, zijn prismatische elementen afgebeeld (15-70 cm), die zijn samengesteld uit kleinere elementen. Bovenin is dit een combinatie van prismatische en blokkige elementen, beneden \pm 30 cm zijnde grote prisma's uit kleinere prisma's opgebouwd.

2. De afmetingen

De grootte wordt voor iedere hoofdvorm in een aantal klassen gescheiden. Deze zijn in fig. 3 eveneens aangegeven met tekens van verschillende grootte.

3. De structuurgraad

De structuurgraad stijgt met het percentage hele structuurelementen dat in de gehele grondmassa door los steken of los breken kan worden gescheiden.

In fig. 3 geeft een dikkere omgrenzing de hele structuurelementen aan, veel dikke omgrensde elementen wijzen een hoge structuurgraad aan.

Met behulp van de hier globaal aangegeven indeling kan iedere horizont met een bepaald structuurtype worden gekarakteriseerd.

Onder het structuurprofiel wordt verstaan de combinatie van elkaar in verticale richting opeenvolgende structuurtypen.

De twee in fig. 3 gegeven voorbeelden geven een visuele voorstelling van twee sterk uiteenlopende structuurprofielen.

Heeft men een groot aantal dergelijke structuurprofielen beschreven, dan kunnen deze complexe eenheden opnieuw in een classificatiesysteem worden gegroepeerd.

Let hierbij op het volgende:

Men kies hierbij:

- eigenschappen (bepaalde, in een beperkt aantal, bepaald volgens aangegeven meetmethoden)

- een classificatie van deze eigenschappen
- de wijze van combinatie van de eigenschappen per element
- de wijze van afgrenzing van de elementen
- men specificceert de classificatie van de onderlinge rangschikking van de elementen.

2.3 Waarom beschrijven we

Er bestaat een nauwe relatie tussen het doel (waarom) en de methode (hoe) en keuze van eigenschappen (wat). Het doel kan zeer algemeen zijn, in de zin van voor algemeen gebruik, d.w.z. voor een groot aantal doeleinden tegelijkertijd geschikt, of meer beperkt, zoals b.v. de geschiktheid van zand als metselzand. In de volgende paragrafen komen we hier nog op terug.

3. CLASSIFICATIE

3.1 Inleiding

In het voorgaande zijn we op zeer veel plaatsen al met aspecten van classificatie geconfronteerd. In deze paragraaf wordt de classificatie, en speciaal het opstellen van classificatiesystemen van complexe elementen, nader besproken.

Classificatie betekent rangschikking. Een classificatiesysteem is een rangschikking van klassen in een systeem. Men deelt een bepaalde groep van objecten in klassen in, op een zodanige manier dat een overzichtelijke ordening wordt bereikt. Er bestaat discussie over de zin en onzin van classificatie, waarbij o.a. twee extreme standpunten worden ingenomen.

De tegenstanders van classificatie stellen dat er maar één goede manier is om objecten aan te duiden, nl. door ordinatie (een ordinaat is één van de beide coördinaten van een punt in een plat vlak). Ieder object wordt door een groot aantal eigenschappen beschreven en per eigenschap geeft men de waarde die deze voor dit object heeft. Dit zou bijvoorbeeld een rij van 50 getallen kunnen zijn. Hierbij geeft men een maximum aan informatie, wat samengaat met een minimum aan overzichtelijkheid.

In tegenstelling tot de ordinatie zal er bij de classificatie verlies aan informatie optreden, maar neemt daarentegen de overzichtelijkheid toe. Of dit zinvol is hangt af van het doel dat men nastreeft.

Twee grote groepen van doeleinden leiden tot natuurlijke resp. kunstmatige classificatiesystemen.

Een natuurlijk classificatiesysteem geeft inzicht in het ontstaan van en de onderlinge relaties tussen de objecten. Een zeer fraai natuurlijk systeem is het Periodiek systeem van de elementen. Dit geeft een ordening die voor een veelzijdige uitleg vatbaar is en die tevens met bouw van de atomen in overeenstemming is. In de bodemkunde zijn er onderzoekers, die betwijfelen of er wel een natuurlijk systeem mogelijk is. Daarnaast zijn er onderzoekers die een bijna mystiek geloof in hun visie op het natuurlijk systeem belijden.

Naast natuurlijke onderscheidt men kunstmatige of technische classificatiesystemen. De eenvoudigste vorm hiervan is de indeling naar één eigenschap (single value), b.v. naar de zwaarte (lutumgehalte) van de bovengrond.

Voor een bepaald gebruik kan dit een zinvolle indeling zijn, b.v. in verband met grondbewerking. De eenzijdigheid geeft daarentegen ook een sterke beperking van de interpretatiemogelijkheden. Complexere kunstmatige indelingen bestaan uit klassen die door meerdere eigenschappen tegelijk worden bepaald, b.v. zwaarte van de bovengrond, zwaarte van de ondergrond en gemiddelde hoogste wintergrondwaterstand.

De doelstelling van deze technische classificatiesystemen is meestal vrij beperkt en goed omschreven. Heeft men een zeer algemene doelstelling, dan kan het meer voor de hand liggen om met een z.g. natuurlijk systeem te werken.

3.2 Eigenschappen van classificatiesystemen in het algemeen

De meeste classificatiesystemen hebben de vorm van een categorieën-systeem (fig. 4).

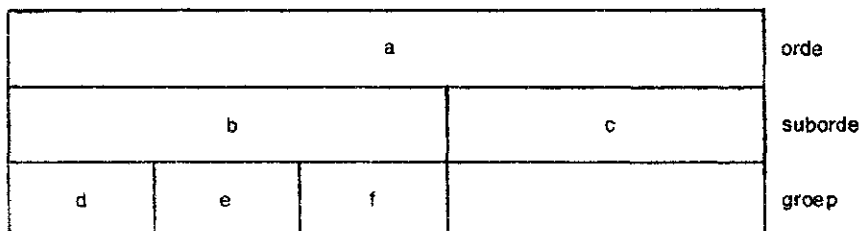


Fig. 4 Principe van een categorieënschema. Orde a wordt onderverdeeld in suborden b en c; suborde b in de groepen d, e en f. Het niveau van b is hoger dan van d, e en f, maar lager dan van a.

Iedere categorie ligt op een bepaald niveau. Klassen worden op het naast-hogere niveau samengevoegd, waarbij een pyramidale opbouw ontstaat. Op het laagste niveau vindt men de grootste hoeveelheid informatie per klasse; hoe hoger men in het systeem komt, hoe geringer de informatie die over een klasse wordt gegeven.

Ieder niveau krijgt een naam, zoals in fig. 4 orde, suborde, etc.

De grenzen van hogere niveaus lopen in het lagere niveau door. Was dit niet het geval dan zouden klassen op verschillend niveau elkaar overlappen en zou men niet tot een eenduidige indeling in het systeem kunnen komen.

Bij het opstellen van een classificatiesysteem kan men in principe tussen twee extreme wegen kiezen:

- "van boven af", of verdelend (divisief) te werk gaan
- "van onder af", of samenvoegend (agglomeratief) te werk gaan.

In de praktijk doet men meestal zowel het ene als het andere, waarbij ervoor moet worden gezorgd, dat de verdeling, die men op een hoger niveau heeft aangebracht, consequent naar beneden toe wordt gehandhaafd.

Wat is een klasse?

Als we ons een beeld willen vormen van een bepaalde klasse, b.v. suborde b in fig. 4, staan er twee wegen open:

- er kan worden uitgegaan van het centrale concept, dat is het ideale beeld van de abstracte meest typische vertegenwoordiger van die klasse
- er kan worden getracht om de grens van de klasse zo nauwkeurig mogelijk aan te geven.

Berust de klasse-indeling op slechts één eigenschap, dan is het aangeven van de grenzen eenvoudig. In een natuurlijk systeem, waarin zeer veel eigenschappen worden gehanteerd, is dit minder eenvoudig.

Stel dat we twee suborden b en c willen onderscheiden en dat we een centraal concept voor b, nl. b^1 , en voor c, nl. c^1 , hebben omschreven. Om nu uit te maken of een nieuw in te delen object x tot b of c behoort bepalen we de mate van gelijkenis met b^1 en met c^1 .

In fig. 5 is aangegeven hoe dit bij het gebruik van twee eigenschappen zou kunnen gebeuren.

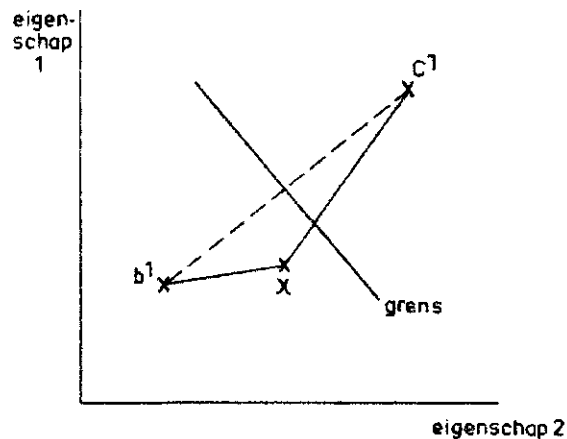


Fig. 5 De toedeling van een object x aan suborden b of c, op basis van de afstand tot de centrale concepten b^1 en c^1 in de tweedimensionale ruimte van de eigenschappen 1 en 2.

De afstand van x tot b^1 is kleiner dan die van x tot c^1 , m.a.w. x komt in eigenschappen meer overeen met b^1 dan met c^1 . Op grond van deze overweging zou x tot de suborde b moeten worden gerekend.

We kunnen ons model nu wat ingewikkelder maken zoals in fig. 6 is weer-
gegeven

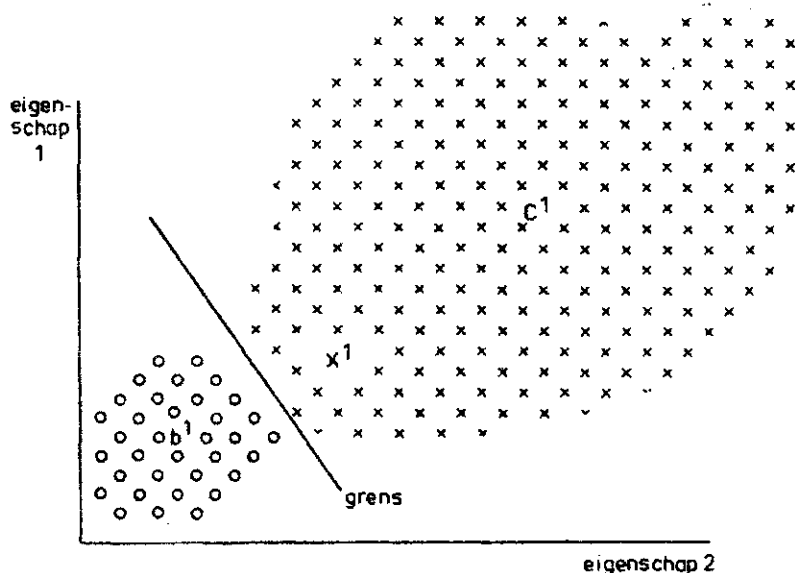


Fig. 6 De toedeling van een object x aan de suborden b of c , waarbij rekening wordt gehouden met de spreiding in eigenschappen en het aantal individuen in b en c .

Met x zijn de individuen in suborde c aangegeven, met o de individuen in suborde b . Wegens de grotere spreiding in eigenschappen en het grotere aantal individuen in suborde c ligt het in dit geval voor de hand om x tot suborde c te rekenen. Wanneer de grens tussen b en c was vastgelegd zou op eenvoudige wijze kunnen worden aangegeven tot welke suborde x behoort.

In moderne bodemclassificatiesystemen tracht men daarom de grenzen te definiëren en hanteert men niet uitsluitend meer de centrale concepten voor het karakteriseren van klassen.

3.3 De eisen die men aan de klassen afzonderlijk kan stellen

Iedere klasse dient zo homogeen mogelijk te zijn. Er is al aangegeven dat met opklimmend niveau de variatie in eigenschappen binnen klassen toeneemt. Op een bepaald niveau zal moeten worden gestreefd naar optimale homogeniteit. Liefst zou de variatie in eigenschappen in iedere klasse op hetzelfde niveau ook ongeveer even groot moeten zijn.

Bij een kunstmatig systeem met twee eigenschappen (1 en 2) zou dit er b.v. als volgt uit kunnen zien (fig. 7).

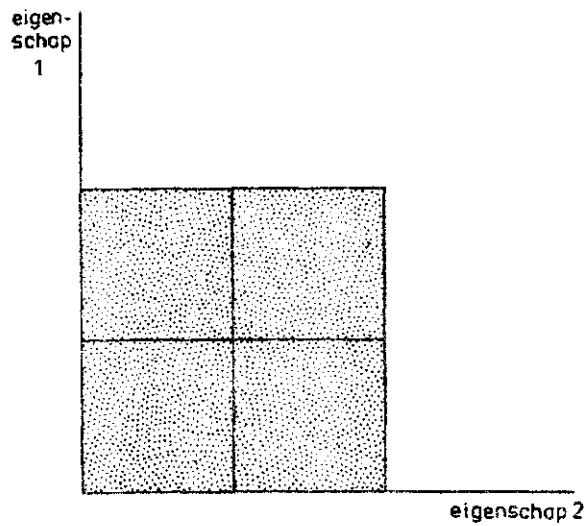


Fig. 7 Een kunstmatige indeling van een universum, waarin de elementen (stippen) gelijkmatig over de eigenschappenruimte zijn verdeeld.

Wanneer de elementen (stippen) gelijkmatig over de ruimte zijn verdeeld, is dit wel een verantwoorde oplossing. In dit geval kan men niet van het voorkomen van natuurlijke grenzen spreken en is voor een natuurlijk systeem de gehele beschouwde eigenschappenruimte één klasse.

In fig. 8 is een geval weergegeven waarbij zeer uitgesproken natuurlijke groepen bestaan, die met een streeplijn zijn omgrensd.

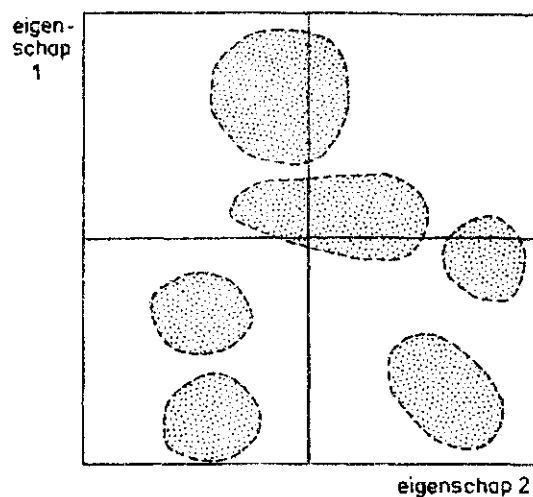


Fig. 8 Een kunstmatige indeling van een universum, waarin de elementen (stippen) in duidelijk te onderscheiden natuurlijke groepen (streeplijnen) zijn samengebond.

Het is in dit geval zonder meer duidelijk hoe een natuurlijke klassenindeling eruit zou zien. De variatie in eigenschappen verschilt per klasse vrij sterk. De gelijkmatige verdeling van de eigenschappenruimte in even grote delen (getrokken lijnen) heeft hier een zeer ongunstig effect.

Wanneer een indeling op een groot aantal eigenschappen van de elementen wordt gebaseerd, maakt men gewoonlijk een scheiding tussen differentiërende kenmerken en begeleidende eigenschappen.

Differentiërende kenmerken zijn de eigenschappen waarop de indeling is gebaseerd. Ze dienen voor de identificatie of het determineren van de klassen, zoals dat b.v. in een flora wordt gedaan.

Begeleidende eigenschappen zijn eigenschappen die met de differentiërende kenmerken zijn gecorreleerd. Hoe groter het aantal begeleidende eigenschappen van een klasse is, en naarmate ze nauwer met de differentiërende kenmerken zijn gecorreleerd, des te beter is de classificatie. Men kan op basis van deze klasse dan een veel groter aantal en nauwkeuriger uitspraken doen, dan op basis van de differentiërende kenmerken alleen.

Over het algemeen kunnen we dus stellen dat er bij een classificatie moet worden gestreefd naar een zo groot mogelijke homogeniteit van de klassen met betrekking tot een zo groot mogelijk aantal relevante eigenschappen.

3.4 Methoden om een classificatiesysteem op te stellen

3.4.1 Algemene randvoorwaarden

Wanneer een classificatiesysteem moet worden opgesteld, dienen een aantal elementaire randvoorwaarden en uitgangspunten eerst duidelijk te worden geformuleerd. Dikwijls gebeurt dat niet uitdrukkelijk. Bij de toepassing van een bepaald systeem zal men zich steeds een duidelijk beeld van het volgende moeten vormen om er zeker van te zijn, dat niet tegen de elementaire regels wordt gezondigd:

1. Er wordt een doelstelling voor het classificatiesysteem geformuleerd. Deze doelstelling kan zeer algemeen zijn, d.w.z. dat men een zo universeel mogelijk systeem wenst op te zetten. In dat geval zal men geneigd zijn voor een "natuurlijk" systeem te kiezen. De doelstelling kan ook veel beperkter

en zeer specifiek zijn, zoals b.v. een bodemindeling geschikt voor de classificatie naar vertrappingsgevoeligheid, in verband met betreding door koeien.

2. Een definitie van de aard van de elementen die in het systeem dienen te worden geordend.

Voor het geval van een bodemclassificatie zal men b.v. moeten aangeven in welke zin men het begrip bodem wil hanteren, in verband met de doelstelling. Bij een genetisch systeem zal men b.v. de bodem beschouwen tot op de diepte waarop veranderingen in de bodem zijn opgetreden door bodenvorming, b.v. door inspoeling.

3. Een omschrijving van het universum waarvoor het systeem geldig moet zijn. Onder universum verstaan we hierbij de verzameling van elementen waarop het systeem betrekking heeft. Dit kan b.v. zijn: de bodem van geheel Nederland. Het komt nogal eens voor dat men classificaties, die voor een bepaald gebied zijn opgesteld, ver buiten dat gebied wil gaan toepassen. Komen beide universa gedeeltelijk overeen, dan kan dit met enig succes gebeuren. Dikwijls echter zal blijken dat wat typisch was in één gebied, b.v. de zwarte aarden in Rusland, elders b.v. in het middenwesten van Noord-Amerika, allerminst als typisch kan gelden.

3.4.2 De traditionele methode voor het vervaardigen van een classificatiesysteem

Methode is in dit verband een wat ongelukkige term, omdat de werkwijze vaak empirisch is, uiterst complex, en gewoonlijk weinig exact omschreven. Nadat aan de eerder genoemde voorwaarden is voldaan (formuleren van de doelstelling, definitie van de elementen en omschrijving van het universum) volgt een procedure die aan de hand van enige schema's slechts zeer globaal kan worden weergegeven.

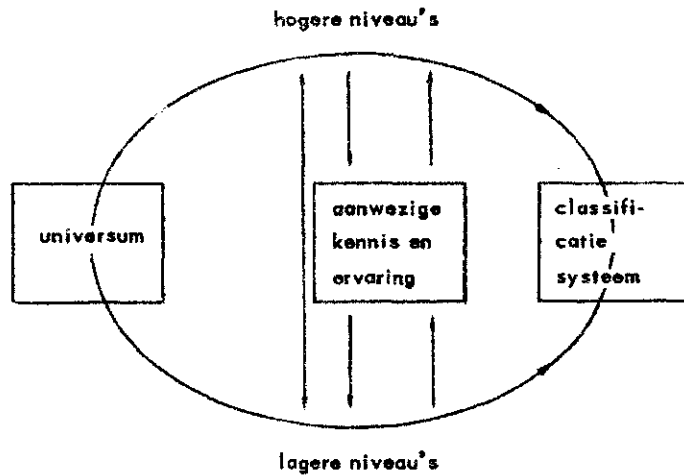


Fig. 9 Algemeen schema van de traditionele methode voor het vervaardigen van een classificatiesysteem.

In fig. 9 is aangegeven dat op grond van monsters uit het universum enerzijds de lagere niveaus van het systeem worden ontwikkeld. Daarnaast werkt men de hogere niveaus van het systeem uit. Er bestaat hierbij een voortdurende wisselwerking tussen deze beide groepen werkzaamheden. Er wordt bovendien in allerlei stadia van beide groepen werkzaamheden steeds getoetst of de groeperingen in overeenstemming zijn met de theorie en ervaring. Dit houdt o.a. in dat men hierbij rekening houdt met bestaande inzichten over het ontstaan (de genese).

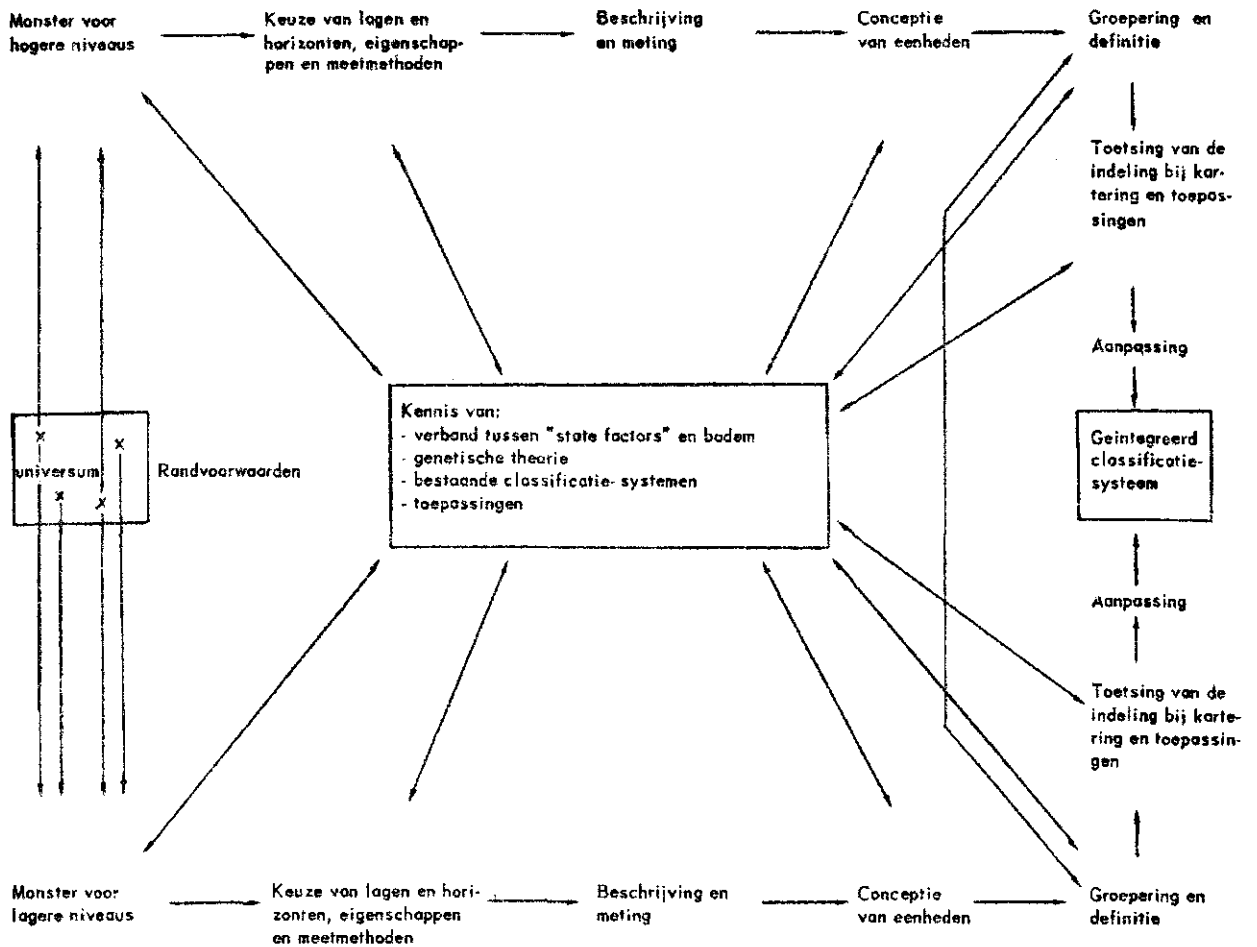


Fig. 10 Schema van de traditionele methode voor het vervaardigen van een bodemclassificatiesysteem.

Schematisch (fig. 10) kan voor beide groepen werkzaamheden voor de hogere en lagere niveaus globaal de volgende procedure worden gevolgd:

1. er wordt uit het universum een monster genomen van elementen, die een goed uitgangspunt vormen;
2. op grond van de doelstelling wordt bepaald in welke onderdelen ieder element kan worden verdeeld (bij de bodem b.v. de lagen en horizonten die in het bodemprofiel kunnen worden onderscheiden). Tevens kiest men die eigenschappen die voor de doelstelling van belang zijn, en legt vast hoe ze worden ingedeeld en gemeten;
3. de bij de bemonstering verzamelde elementen worden met de sub 2 gekozen middelen beschreven en gemeten;

4. op grond van het zo verkregen materiaal en de theorie en ervaring formuleert men nu de concepties van de eenheden;
5. een groepering van deze concepties en een nadere zo exact mogelijke definitie (hetzij van de centrale concepten of van de grenzen) van de eenheden wordt opgesteld;
6. vervolgens toetst men het verkregen resultaat aan de doelstelling en probeert hem uit in de praktijk;
7. op grond van de hierbij verkregen inzichten en ervaringen wordt de indeling en karakterisering vervolgens gewijzigd en aangepast.

Deze sterk geschematiseerde voorstelling geeft slechts een ruw beeld van de werkelijke gang van zaken. Er treden in de praktijk tal van terugkoppelingen op en bepaalde onderdelen worden herhaaldelijk opnieuw doorlopen.

Uit het voorgaande blijkt wel, dat de traditionele manier om een classificatiesysteem op te stellen in veel opzichten ondoorzichtig is. Het is dan ook begrijpelijk dat er naar wegen is gezocht om betere methoden te ontwikkelen. De nieuwe wetenschap, die hierop is gericht, wordt aangeduid als numerieke classificatie of numerieke taxonomie.

3.4.3 Enige aspecten van de numerieke taxonomie

De doelstelling van de numerieke taxonomie is de ontwikkeling van exacte, reproduceerbare methoden voor het opstellen van classificatiesystemen. Iedereen die op hetzelfde materiaal een bepaalde methode toepast krijgt daaruit een identiek resultaat. Dit is op zich zelf een grote vooruitgang ten opzichte van de klassieke werkwijze. Het behoeft nog niet in te houden dat het resultaat ook zinvol is in verband met de doelstelling. Een van de grote problemen, waarvan de oplossing nog urgenter is geworden dan ze al was, is: hoe beoordelen we of een classificatie aan de doelstelling beantwoordt? We kunnen nl. allerlei onderdelen van de methode variëren en zo een gehele reeks verschillende classificaties voor hetzelfde materiaal berekenen.

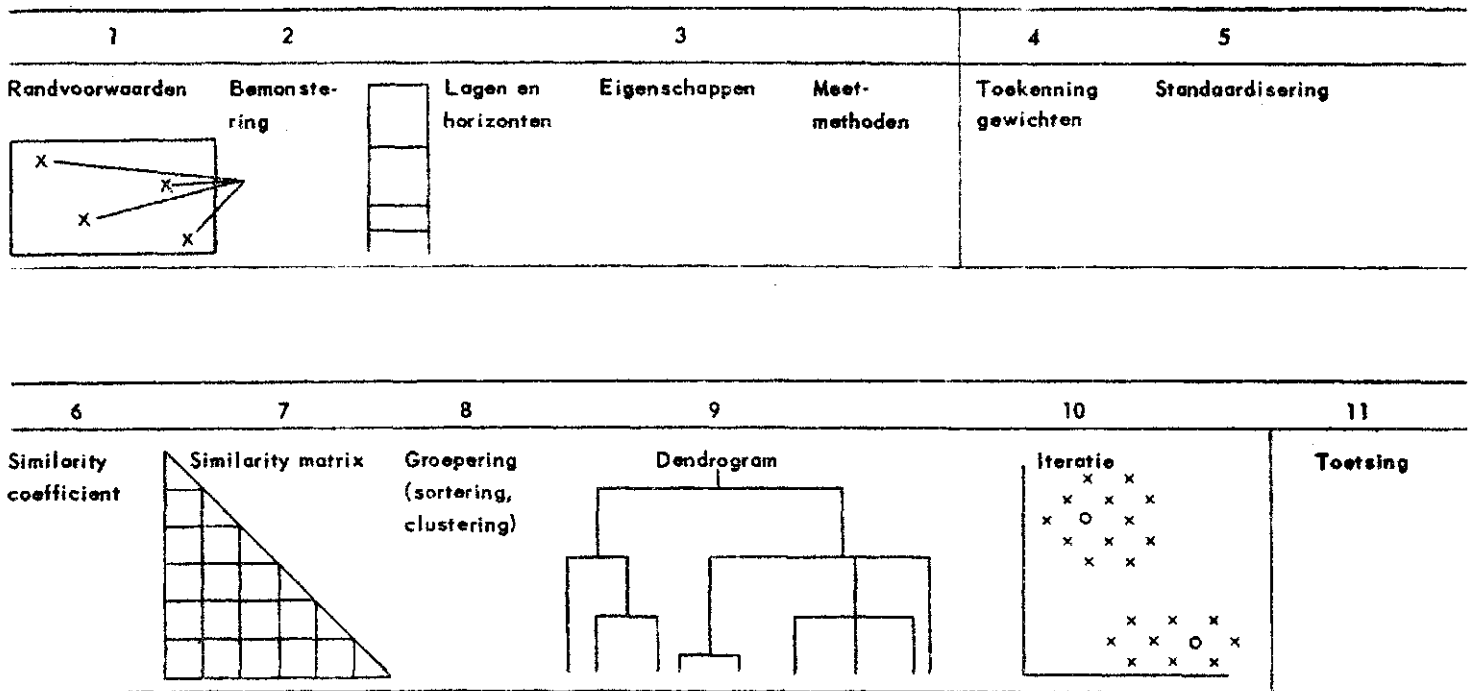


Fig. 11 Schema van het vervaardigen van een bodemclassificatiesysteem met behulp van numerieke methoden.

In fig. 11 is de gang van zaken schematisch weergegeven. De voorbereiding (1, 2 en 3) is identiek aan die van de traditionele methode. Dan volgt het specifieke deel van de numerieke taxonomie. De laatste fase van het testen is weer identiek aan de traditionele werkwijze.

Drie belangrijke onderdelen worden nader toegelicht:

- de mate van gelijkenis
- de groepering of clustering
- de iteratieve methode.

De mate van gelijkenis

De uitspraak A gelijk op B is nog onduidelijk. Pas wanneer er wordt gesteld "A gelijk op B in dit opzicht, of met betrekking tot deze eigenschap (pen)" krijgt deze uitspraak meer betekenis.

Drie gronden, A, B en C, met kleigehalten in de bovenste 20 cm van resp. 5, 20 en 25 % (fig. 12), worden met elkaar vergeleken.

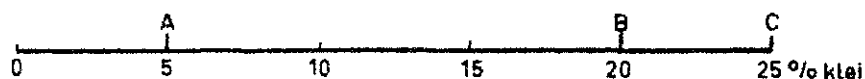


Fig. 12 Het kleigehalte in de bovenste 20 cm van het bodemprofiel als basis voor een indeling. De gronden B en C lijken meer op elkaar dan op A, omdat hun onderlinge afstand op de eigenschapschaal kleiner is.

Het is voor iedereen duidelijk dat B meer lijkt op C dan op A. Dit berust op het verschil in de eigenschap "kleigehalte van de bovenste 20 cm". In een schema kan dit als volgt worden weergegeven (fig. 13).

Aangezien de tabel symmetrisch is kan het gedeelte boven de diagonale lijn worden weggelaten.

	A	B	C
A	0	15	20
B	15	0	5
C	20	5	0

Fig. 13 Matrix van de afstanden tussen alle paren gronden in fig. 12. De matrix is symmetrisch, zodat slechts het gedeelte onder de diagonaal hoeft te worden weergegeven (+ en - tekens zijn niet aangegeven).

Door meerdere eigenschappen te gebruiken kan er meer inhoud worden gegeven aan uitspraken over de mate van onderlinge gelijkheid.

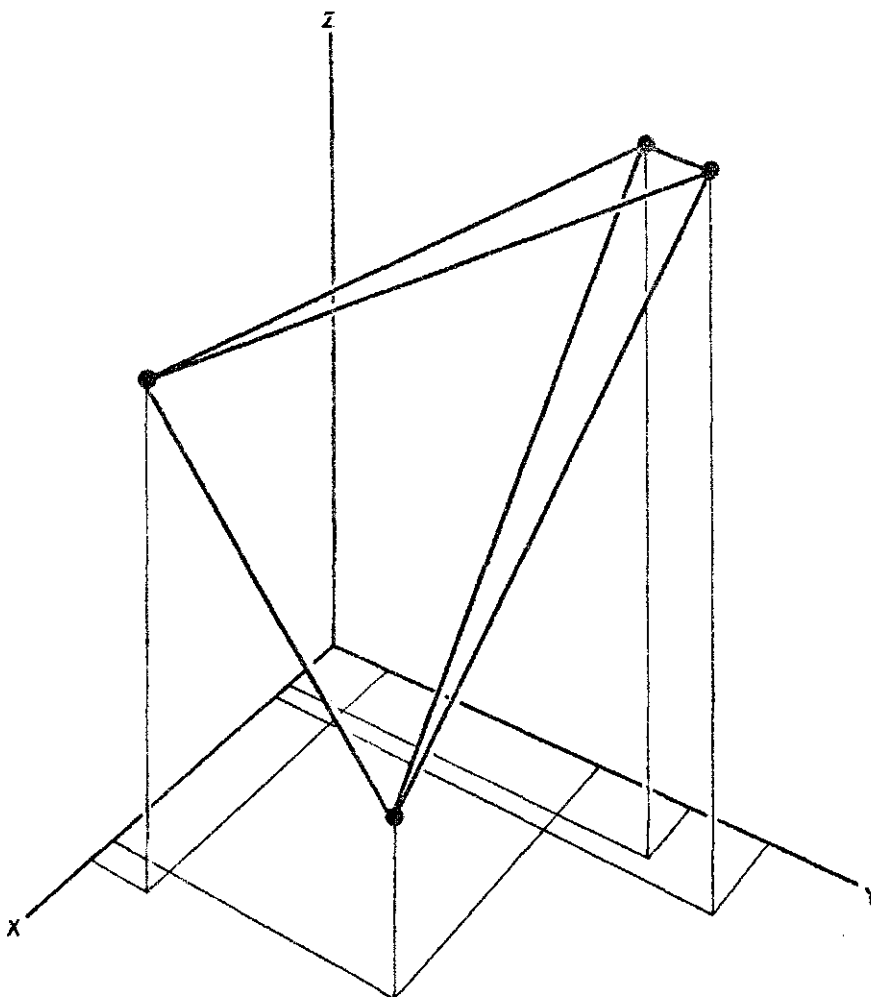


Fig. 14 Vier objecten (zwarte ballen) in een driedimensionale ruimte van de eigenschappen X, Y en Z. Kortere afstand tussen twee objecten houdt een grotere mate van gelijkheid in.

In fig. 14 is op grond van drie eigenschappen X, Y en Z de afstand tussen vier elementen aangegeven.

Het is duidelijk dat de elementen met de kortste afstand het meest op elkaar gelijk zijn.

Ook in een n-dimensionale eigenschappenruimte, waarbij n bij voorkeur groot is (b.v. 30 of 50) kunnen de onderlinge afstanden tussen alle paren elementen worden berekend. Ook hier geldt: hoe korter de afstand, des te meer lijken de elementen op elkaar. Dit levert een gelijksoortige tabel op als bij het lineaire voorbeeld.

Afstand kunnen we beschouwen als een maat voor de ongelijkheid, hoe groter de afstand, hoe meer ongelijk twee elementen zijn. Door het systeem zo op te zetten dat alle afstanden variëren tussen 0 en 1, kunnen we (1 - de afstand AB) als maat voor de gelijkheid gebruiken.

Naast de afstand bestaan er tal van andere coëfficiënten voor de mate van gelijkheid (similarity coefficient), ieder met specifieke eigenschappen, en voor- en nadelen voor bepaalde toepassingen.

De groepering of clustering

Zoals bij de traditionele methode van "boven af" of van "onder af" kan worden gewerkt, bestaan er verdelende (divisie) en samenvoegende (agglomeratieve) methoden in de numerieke taxonomie.

Bij de verdelende methode worden eerst alle mogelijke verdelingen van de totale populatie van elementen in twee groepen vergeleken volgens een of ander criterium, dat de homogeniteit van de groepen aangeeft. De groepering die tot maximale homogeniteit leidt wordt uitgekozen en hetzelfde proces wordt herhaald voor de successievelijke opsplitsing van de gevormde groepen. Over het algemeen vereist dit zeer veel rekentijd en kan men slechts populaties van beperkte omvang op deze wijze bewerken.

De samenvoegende of agglomeratieve methoden komen in grote aantallen voor en kunnen voor grotere populaties worden gebruikt. Met twee voorbeelden ontleend aan Sokal wordt de werking van enige methoden toegelicht.

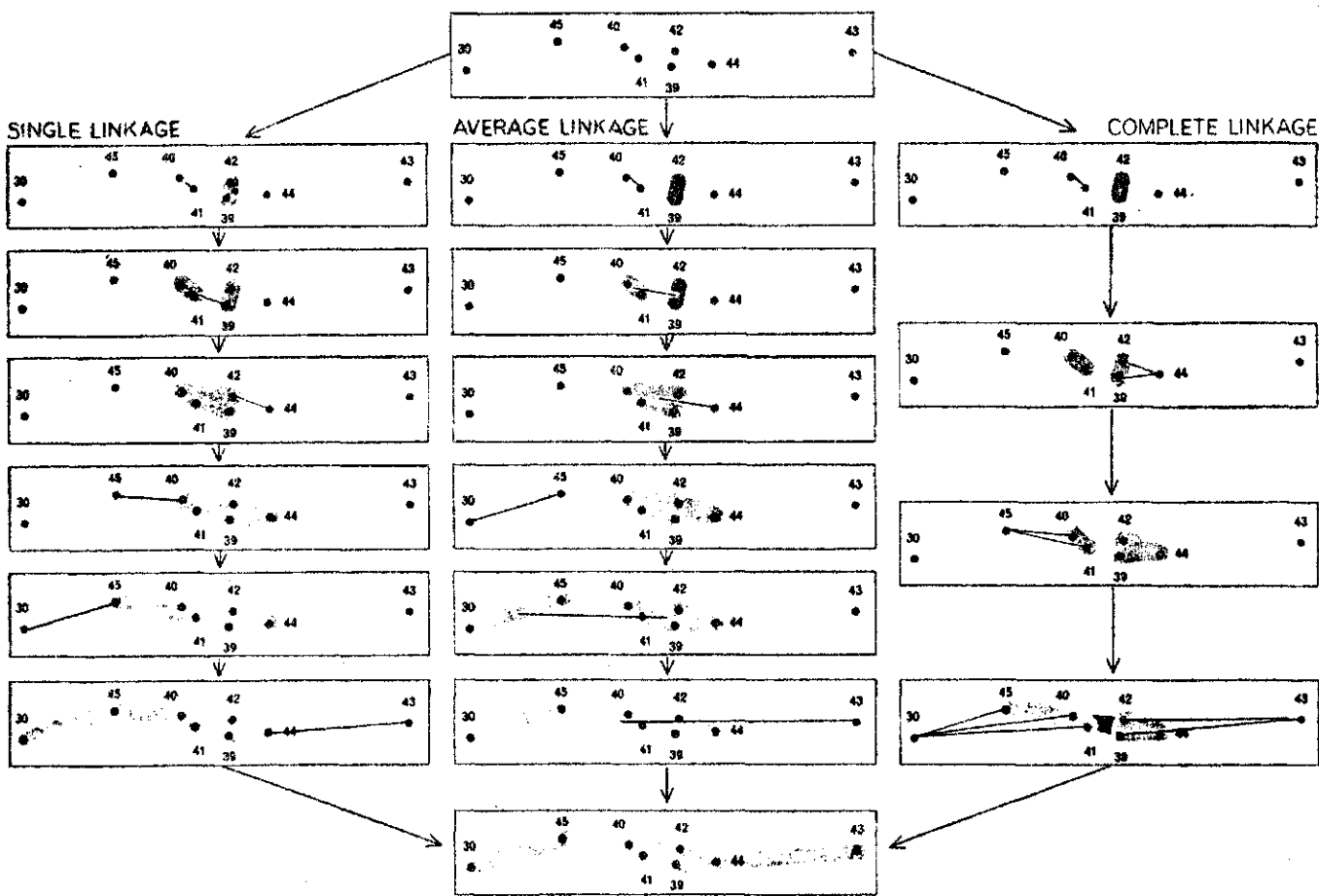
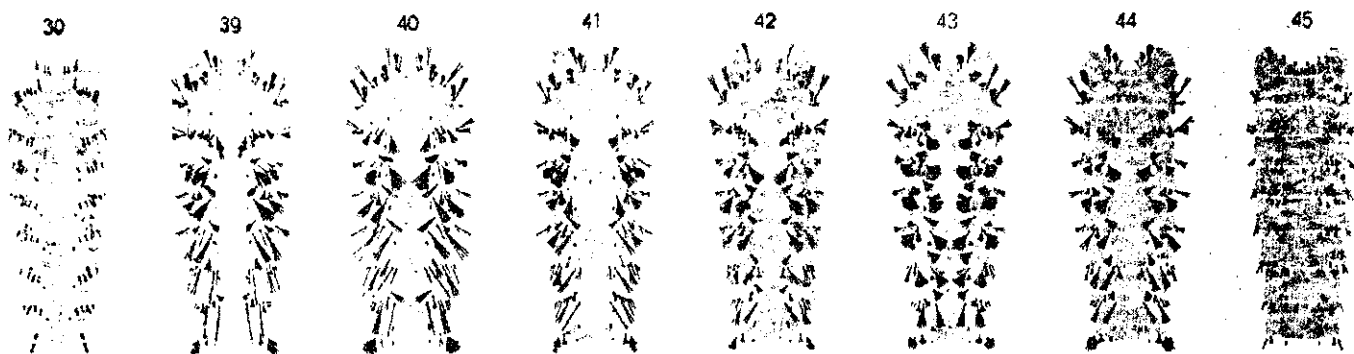


Fig. 15 Drie methoden van groepering (clustering) leveren verschillende classificatiesystemen op voor hetzelfde materiaal. De eigenschappen van de poppen van acht verschillende muskieten zijn in het bovenste rechthoekje in een tweedimensionale ruimte afgebeeld. Nieuwe combinaties zijn door getrokken lijnen, reeds in voorgaande stappen samengevoegde punten in grijs aangegeven.

In fig. 15 wordt de werking van drie methoden, toegepast op hetzelfde materiaal, gedemonstreerd. De genummerde zwarte stippen geven acht soorten muskieten weer, beschreven op grond van de eigenschappen van hun poppen (teke-

ningen boven aan de figuur). Soort 30 behoort tot het genus *Mansonia*, de andere zijn soorten van *Anopheles*. Ter wille van de eenvoud zijn de afstanden tussen de punten slechts in een tweedimensionale ruimte weergegeven. Verbindingen tussen elementen en groepen, en tussen paren groepen, zijn weergegeven door getrokken lijnen. In een voorgaand stadium gevormde groepen zijn met een lichte punctering aangeduid. Naarmate de groepering stap voor stap voortschrijdt worden de criteria voor de samenvoeging minder streng, met andere woorden de afstanden tussen de toegevoegde elementen en de bestaande groepen worden steeds groter.

Bij de "single linkage" methode (links), wordt begonnen met het paar met de geringste onderlinge afstand, en worden successievelijk elementen met toenemende afstand tot één element van de bestaande groep toegevoegd.

Bij de "average linkage" methode (midden) wordt dat element aan een groep toegevoegd dat de geringste afstand heeft tot het zwaartepunt van die groep.

Bij de "complete linkage" methode (rechts) wordt een element aan een groep toegevoegd, wanneer de relatie tussen dat element en alle leden van de groep minimaal is voor die fase in de groeperingscyclus.

Hoewel het beginpunt en het eindpunt gelijk is voor alle drie methoden, zijn de tussenliggende stappen meestal verschillend in de verschillende stadia.

Een ander voorbeeld van de toepassing van numerieke taxonomie is weergegeven in de fig. 16 en 17.

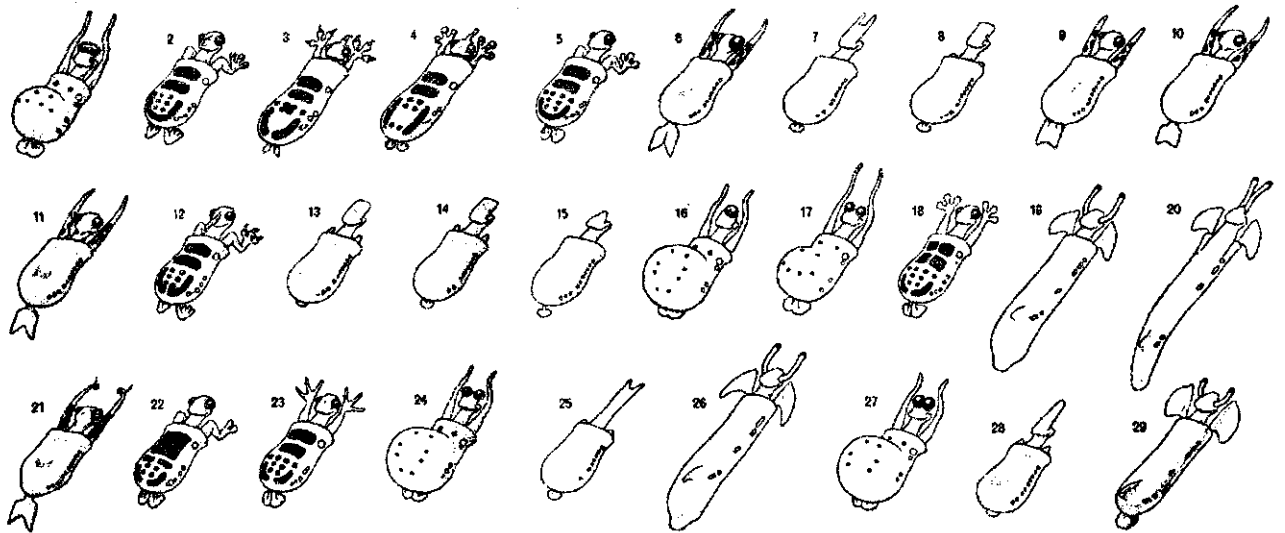


Fig. 16 Fantasiebeesten Caminalcules genaamd naar hun schepper J.H. Camin. Deze z.g. "recente" soorten zijn ontstaan als gevolg van een door Camin bedachte evolutie.

In fig. 16 zijn fantasiebeesten afgebeeld, Caminalcules genoemd naar hun schepper Joseph H. Camin van de Universiteit van Kansas. Deze z.g. "recente" dieren zijn afgeleid volgens een door hem opgesteld schema van evolutie. In fig. 17 is het resultaat van een met numerieke methoden berekende classificatie weergegeven. Het blijkt dat het resultaat goed overeenstemt met de visuele indruk van de mate van gelijkheid.

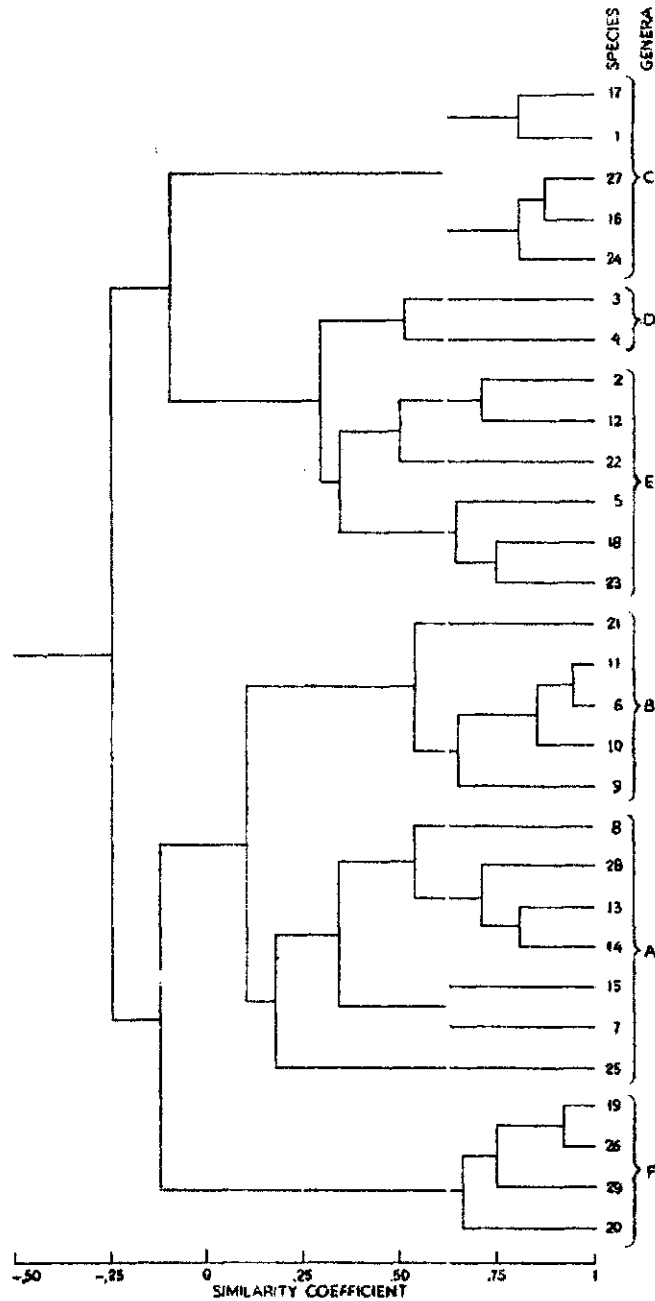


Fig. 17 Dendrogram, schematische weergave van een classificatiesysteem, dat met numerieke methoden is berekend voor de z.g. soorten Caminalcules van fig. 16.

De soorten species zijn met nummers aangegeven aan het einde van de takken, de klassen (genera) met hoofdletters.

De weergegeven genera A, B, etc. zijn niet geheel consequent samengesteld. Ze behoren op één zelfde niveau van de mate van gelijkheid (similarity coefficient) te worden gescheiden, waarbij soort 25 tot een aparte klasse van dieren zou behoren (trek hiervoor een verticale lijn in het schema bij 0.32).

De iteratieve methode

Wanneer voor een monster uit het universum (een aantal elementen, b.v. bodemprofielen) een dendrogram berekend is, en hieruit is een keuze gedaan van het niveau dat men als het laagste wenst te hanteren, dan staat hierdoor het aantal groepen vast. Voor iedere groep kan nu het centroid worden berekend, het zwaartepunt in de n-dimensionale eigenschappenruimte.

Nemen we nu een nieuw monster uit het universum, dan rijst de vraag hoe we deze nieuwe elementen aan de gevormde groepen kunnen toedelen en of dit resultaat nog verbeterd kan worden.

Voor ieder nieuw element kunnen de afstanden tot alle centroiden worden berekend. Het element wordt dan gevoegd bij de groep waartoe het de kortste afstand heeft. Is dit voor alle nieuwe elementen gedaan, dan kan men voor iedere nieuwe groep een nieuw centroid berekenen, dat meestal verschoven is ten opzichte van het voorgaande. Dit kan net zo lang worden herhaald tot de centroiden niet meer verschuiven.

Het gehele proces leidt weliswaar niet tot een absoluut optimum en is b.v. afhankelijk van de ligging van het startpunt.

3.4.4 Voorbeeld van een classificatiesysteem

Als voorbeeld voor een op traditionele wijze vervaardigd classificatiesysteem is het "Systeem van bodemclassificatie voor Nederland" gekozen. In de brochure "Indeling van Nederlandse gronden" (J. Schelling, H. de Bakker en G.C.L. Steur, 1967) is in het derde deel (na pag. 15) een globale omschrijving gegeven.