

# Samenvatting

**Hengl, T. 2003. Pedometrische kartering: overbrugging de kloof tussen de traditionele- en pedometrische benaderingen. Doctoraal proefschrift, Universiteit Wageningen.**

Digitale bodemkartering heeft de afgelopen jaren een snelle ontwikkeling doorgemaakt door nieuwe- en economische methoden, als gevolg van het beschikbaar komen van vele soorten secundaire kaarten (terrein parameters, satellitbeelden etc.). Het hoofddoel van dit onderzoek was om een methodologie te ontwikkelen voor pedometrische kartering die gebruikt kan worden om de kloof te dichten tussen de mechanistische pedometrische- en de traditionele benaderingen. De thesis beslaat zeven methodologische benaderingen van bodemkartering: bemonstering, voorbereiding, foto-interpretatie, interpolatie, visualisatie, organisatie en kwaliteitscontrole. Deze hoofdstukken zijn ingediend als wetenschappelijke artikelen in internationale tijdschriften.

**BEMONSTERING:** Dit hoofdstuk behandelt de verspreiding van waarnemingen in de data- en geografische ruimte als sleutel tot de optimalisatie van bemonstering voor de ruimtelijke voorspelling tijdens de correlatie met secundaire kaarten. Ofschoon deze kaarten veel gebruikt worden voor de kartering van bodemvariabelen worden de problemen die te maken hebben met het ontwerp van een bemonsterings schema zelden onderzocht. Wanneer de gegeneraliseerde kleinste kwadraten wordt toegepast hangt de totale voorspellingsfout af van de verspreiding van de observatie punten over de data- en de geografische ruimte. Het verdelen van bemonsteringspunten over de dataruimte in proportie tot de verdeling van de voorspeller (*equal range* stratificatie of ER ontwerp) wordt voorgesteld als een verstandige bemonsterings strategie wanneer het regressie model tussen de bodem en de ondersteunende variabelen niet bekend is. Om dit concept te illustreren werd een gebied van 50×50 km in centraal Kroatië met 100 waarnemingen gebruikt. Het gebied werd herbemonsterd tot 25 databeelden met verschillende experimentele ontwerpen: ER en twee responsie-oppervlak ontwerpen (minmax en D2). Het ER gaf dezelfde voorspellingsfout als het minmax ontwerp, wat aangeeft dat het een goed compromis is tussen nauwkeurige schatting van het model en de minimalisatie van de ruimtelijke autocorrelatie van de restwaarden. Ook werd het ER ontwerp uitgebreid naar een multi-variatie studie. Vier voorspellers (hoogte, temperatuur, vochtigheidsindex en NDVI) werden getransformeerd tot hoofdcomponenten. De bemonsteringspunten werden vervolgens toegewezen aan

de componenten in verhouding tot de variantie die bepaald werd door de Hoofdcomponenten Analyse en volgens het ER ontwerp.

**VOORBEWERKING:** De kwaliteit van de toepassingen van de terrein analyse hangt direct af van de kwaliteit van de digitale hoogte modellen (DEMs) en de hiervan afgeleide data. Drie benaderingen voor reductie van fouten in DEMs en afgeleide data werden beschreven: (a) het gebruik van empirische kennis, b.v. aanpassing van de hoogte d.m.v. mediale assen en het drainage netwerk; (b) data filtreren en (c) foutenvoortplanting. Data filtreren werd gebruikt om onregelmatige waarden te vervangen of om uitschieters te reduceren door middel van de ruimtelijke correlatie structuur en de waarschijnlijkheid van het overschrijden van een waarde zoals geschat door naburige waarden. Bij de foutenvoortplanting benadering worden de fouten teruggebracht door berekening van de waarden van verschillende realisaties. Deze methode werd getest in een proefgebied van  $3.8 \times 3.8$  km waarin twee landschappen voorkomen: heuvel-land en een vlakte met terrassen. Het gedeelte aan artefacten (padi-terrassen) in de onbehandelde DEM was 17.3%. Na toevoeging van mediale assen, het filtreren van uitschieters, en de aanpassing van de hoogten van het drainagepatroon, werd het aandeel van de padi-terrassen teruggebracht tot 2.2%. Uitschieters bestreken over het gehele gebied niet meer dan 2%. Zowel filtreren als de foutenvoortplanting benadering resulteerden in een wat gelijkmatiger kaartbeeld van de terrein parameters. Het voordeel van filtering van uitschieters is dat het gebruik maakt van de structuur van de ruimtelijke afhankelijkheid. Het voordeel van de foutenvoortplanting benadering is dat het gemakkelijker geautomatiseerd kan worden. Vermindering van de fouten verbeterde de kartering van de landschap facetten (classificatie) en van de dikte van de solum (regressie). De nauwkeurigheid van de classificatie nam toe van 51.3% tot 72% en de  $R^2$  van het regressie model voor de voorspelling van de solumdikte verbeterde van 0.27 tot 0.40.

**FOTO-INTERPRETATIE:** Een methode wordt beschreven die de handmatige omlijn-ning van landvormen door middel van foto-interpretatie verhoogt voor het karteren van een groter gebied. Traditionele API- kaarten met een geo-pedologische legenda van 21 klassen werden vervaardigd voor zes testgebieden met een totale oppervlakte van 11 sq.km in de Baranja regio van oost Kroatië. Negen terrein parameters werden ontleend aan een digitaal hoogte model, (diepte tot het grondwater, terrein kromming, karakter van de helling, *viewshed*, stroomgebied, vochtigheids index, sediment transport index, en de afstand tot de dichtstbijzijnde waterloop). Deze werden gebruikt om de luchtfoto-interpretatie te extrapoleren over het gehele gebied (1062 sq.km). De nauwkeurigheid van de classificatie werd vastgesteld door de fouten matrix die berekend werd door de vergelijking van de hele API kaarten en bemonsteringspunten met de resultaten van de classificatie. De eerste resultaten berekend door een hoogste aan-nemelijkheid (*maximum-likelihood*) classificatie waren een reproduceerbaarheid van 58.2% voor heuvelland, 39.1% voor vlakte en 45.3% voor het gehele gebied. Zes klassen in de vlakte waren verantwoordelijk voor een groot aandeel foute classificaties. Dit was het gevolg van een onvoldoende gedetailleerd digitaal hoogte-model en het complexe karakter van de landvormen (kronkelwaarden en actieve rivierlopen), die niet verklaard kunnen worden door alleen gebruik te maken van de terrein parameters.

Door gebruik te maken van een vereenvoudigde legenda van 15 klassen voor het onderzoeksgebied verbeterde de reproduceerbaarheid tot 65.8% voor vlakke, 58.2% voor heuvelland en tot 63.4% voor het gehele gebied. Na vereenvoudiging van de legenda tot 15 klassen en door gebruik te maken van de iteratieve selectie (3 stappen) van het oefenbestand van de bemonsteringspunten werd de classificatie verhoogd tot 97.6% heuvelland, 86.7% vlakke en 90.2% voor het hele gebied. De afgeleide classificatie was in staat om kleine details te laten zien die niet konden worden waargenomen m.b.v. luchtfoto-interpretatie. Het aantal handmatige luchtfoto-interpretaties werd teruggebracht van 84 naar 6. De methodologie kan worden toegepast door bodemkarterers om bestaande kaarten te verfijnen en voor verbetering of vervanging van de luchtfoto-interpretatie voor nieuwe kartering.

**INTERPOLATIE:** In dit hoofdstuk wordt een methodologische benadering beschreven voor ruimtelijke voorspelling gebaseerd op regressie-kriging, en wordt deze vergeleken met gewone kriging en normale regressie. De gegevens werden eerst omgezet door logit transformatie voor doelvariabelen en factor analyse voor de continue voorspellers (secondaire kaarten). De doel variabelen werden vervolgens ingepast door middel van stapsgewijze regressie en restwaarden werden geïnterpoleerd door middel van kriging. Een generieke visualisatie methode wordt gebruikt om gelijktijdig de voorspellingen en de hiermee geassocieerde onzekerheid te visualiseren. Deze benadering werd getest voor 135 waarnemingspunten van de nationale gegevensbank van Kroatië en verdeeld in 100 interpolatie- en 35 validatie punten. Drie doel variabelen (organische stof, pH van de bovengrond en dikte van de bovengrond) werden voorspeld op basis van zes relief parameters en negen bodemkarteringseenheden. De voorspellingen efficiëntie werd geëvalueerd door gebruik te maken van de gemiddelde fout en de gemiddelde gekwadrateerde fout op de validatie punten. De resultaten tonen aan dat de voorgestelde structuur de efficiëntie van de voorspellingen verbeterd. Bovendien garandeert het de normaliteit van de restwaarden en dwong het de voorspelbare waarden binnen de fysieke grenzen van de variabele te blijven. Voor organische stof werd een lagere wortel van de gestandaardiseerde gemiddelde gekwadrateerde voorspelfout bereikt dan met gewone kriging (53.3% versus 66.5%). Dit was ook het geval voor de dikte van de bovengrond (66.5% versus 83.3%), waar ook een kleinere afwijking werd gevonden dan met met gewone kriging (0.15 cm versus 0.69 cm). De voorspelbaarheid van de pH was moeilijk met alle drie methoden. Deze benadering biedt de mogelijkheid om een samengesteld algoritme te ontwikkelen dat kan worden toegepast in een GIS om gegevens van bodemobservatie punten te interpoleren van bestaande databestanden.

**VISUALISATIE:** Een methode om meervoudige lidmaatschappen (multiple memberships) kaarten te visualiseren, nl. kleuren meng (*Colour Mixture* of CM) methode, wordt beschreven, en vergeleken met twee alternatieve technieken: defuzzifikatie en pixel menging (*Pixel Mixture*). Zes landvorm parameters werden gebruikt om de de landvorm klassen te definiëren met gebruik making van de *fuzzy k-means* classificatie. De continue categorische kaart is gebaseerd op GIS berekeningen met kleuren, waarbij de kleurwaarden (Hue) verondersteld worden om de taxonomische ruimte van de variabelen te representeren. Eerst werden de centrale punten van de negen landvormen getransformeerd van de multivariate naar de twee-dimensionale data-ruimte door factor

analyse en vervolgens geprojecteerd op het nl. Kleurwaarde-Verzadiging-Intensiviteit (*Hue Saturation Intensity* of HSI) kleurenschema. De taxonomische waarde werd uitgedrukt met de kleurwaarde en de verwarring met verzadiging. De laatste waarde werd vervangen door 'witheid' om de visualisatie te verbeteren. Klassen die dichter bij elkaar lagen in de data ruimte werden samengevoegd in gelijkwaardige generieke kleuren. De CM techniek beperkt de afgeleide gemengde kleurenkaart tot zeven generieke kleurwaarden, onafhankelijk van het totaal aantal klassen, en vormt de basis voor de automatische generalisatie. De index voor verwarring (*confusion index*) die werd afgeleid van de gemengde kleurenkaart werd gebruikt om primaire grenzen te bepalen en om gebieden met hogere taxonomische onnauwkeurigheid te lokaliseren.

**ORGANISATIE:** In dit hoofdstuk worden basis concepten, handelingen en de structuur van een raster gebaseerd bodem-informatie systeem vergeleken met een traditioneel polygon gebaseerd bodeminformatie systeem (*Soil Information System* of SIS). Dit wordt geïllustreerd door een studie van een gebied van 3.8×3.8 km in oostelijk Kroatië. De ruimtelijke eenheid in dit systeem is een raster cel en alle GIS kaarten werden teruggebracht tot dezelfde resolutie, namelijk 25 m. Bodem-variabelen werden gemodelleerd met behulp van een gemengd model van ruimtelijke variatie zodat zowel discrete als continue variatie mogelijk was. Het bodeminformatie systeem (SIS) bevatte 21 bestaande kaarten (luchtfoto interpretatie, terrain parameters en satelliet beelden), alsmede zes kaarten van bodem variabelen (bodemdikte, aanwezigheid van een mollic, calcic of gleyic horizon, dikte van de bovengrond en percentage silt van deze laag), alsmede zes afgeleide kaarten van bodem typen. Iedere bodem-variabele werd geïnterpoleerd door middel van regressie-kriging. De geïnterpoleerde kaarten werden vervolgens geclassificeerd door gebruik te maken van een continue classifier (*fuzzy k-means*) om lidmaatschapskaarten te vervaardigen. Deze kaarten werden vervolgens gebruikt om de geschiktheid voor tarwe productie te evalueren op een schaal van 0–1. Het bleek dat de luchtfoto-interpretatie kaart beter in staat was om de aangegeven bodemvariabelen te voorspellen dan de terrein parameters en satellietbeelden. Een vergelijking gebaseerd op de aanpassingsmaat en de thematische verwarring toonde aan dat het raster gebaseerde SIS over het algemeen beter overeenkomt, met meer detail geeft en een hogere mate van betrouwbaarheid heeft dan het traditionele polygon gebaseerde SIS.

**KWALITEITSCONTROLE:** Een methodologie om de kwaliteit en de geschiktheid van een nationale bodeminventarisatie te bepalen en om haar bruikbaarheid te evalueren, worden beschreven in dit hoofdstuk. Daarvoor werden gebruikt: zes kaartbladen schaal 1:50 K (uit een totaal van 180), drie proefopnamen (elk met een grootte van 4×4 km) en tien complete profielbeschrijvingen van de hoofdlandschappelijke regio's van Kroatië. Deze gegevens werden gebruikt om de effectieve kaartschaal te bepalen, de nauwkeurigheid van de kaartlegenda's te controleren en de thematische nauwkeurigheid van de bodemobservatie punten te evalueren. Ook werden bestaande digitale bodeminformatie bestanden geevalueerd op hun thematische zuiverheid en op kleigehalte, pH en organische stof getoetst. Nieuwe methoden werden ontwikkeld en getest om de ruimtelijke nauwkeurigheid van de bodemgrenzen en de thematische overlap tussen kaartenheden te onderzoeken. In het proefgebied kwamen de gemid-

delde grootte van de polygon en de positie nauwkeurigheid van de bodemgrenzen (ongeveer  $\pm 40$  m) overeen met de 1:150 K kaartschaal, terwijl de correlatie dichtheid correspondeerde met de 1:250 K kaartschaal. Kaarteenheden waren heterogeen met een gemiddelde relatieve variatie van 17% binnen de eenheden en een gemiddelde thematische overlap van 66% tussen geografisch aangrenzende eenheden. Er is een groot verschil tussen de originele legenda en de proefbemonstering ter validatie als het gaat om de taxonomische classificatie, maar dit is minder het geval als de klassen worden gegroepeerd op basis van overeenkomstigheid. De voornaamste gebruiksproblemen betroffen samengestelde kaarteenheden, gebrek aan specifieke interpretaties die overeenkwamen met de behoeften van de gebruiker en het gebrek aan duidelijk legale wetgeving op het gebied van landeigendom en het -gebruik.

De algemene conclusie is dat de voorgestelde pedometrische kaarterings methodologie bodemkartering verbetert omdat het de bodemkaarten meer objectief, gedetailleerder en meer vergelijkbaar maakt voor integratie met andere ruimtelijke geodata. Er is verder geen noodzaak om nog gebruik te maken van het concept van kaarteenheden of de polygon-gebaseerde bodemkaarten. Daar tegenover staat dat inplaats van het afschaffen van de luchtfoto-interpretatie, de bodem classificatie en de empirische kennis aangaande de bodem, deze methoden succesvol geïntegreerd kunnen worden met pedometrische technieken.