

Het belang van een goede bemonstering en monitoring

Kwantitatieve analyse van ruimtelijke gegevens wordt steeds belangrijker bij maatregelen ter bescherming van de bodem. De noodzaak wordt steeds duidelijker om op basis van een beperkt aantal waarnemingen uitspraken te doen in termen van "het terrein is met 98% zekerheid niet verontreinigd" en "het volume van de grond waarvan de concentratie aan een vervuilend element boven de C-waarde ligt wordt door deze contour aangegeven" en "om een goede inventarisatie te krijgen zijn 160 waarnemingen nodig". Bij het maken van kostencalculaties speelt de onzekerheid in de bepaling van het volume verontreinigde grond een belangrijke rol. Na een sanering moet gecontroleerd worden of de gesaneerde grond inderdaad schoon is en schoon blijft. Daarnaast moet men een aantal potentiële risicolocaties intensief (blijven) monitoren om te zien of aan bestaande normen wordt voldaan. Daardoor bestaat er toenemende belangstelling voor een wetenschappelijke onderbouwing van bemonsterings- en monitoring activiteiten. Telkens terugkerende vragen zijn het aantal monitoringspunten, de localisatie van deze punten en de frequentie van monitoring. Voor de verwerking en analyse van monitoring gegevens kunnen geostatistische methoden toegepast om de snelheid van veranderingen en de oorsprong van een verontreiniging in ruimte en tijd te bepalen en om een interpretatie van de ruimte-tijds dynamiek te geven. De meeste aandacht is tot op heden uitgegaan naar vragen van ruimtelijke aard. De belangstelling voor ruimte-tijds interacties begint steeds groter te worden.

Alfred Stein en Johan Bouma

Monitoring en variabiliteit in ruimte en tijd

Bij monitoring worden waarnemingen verricht op verschillende locaties, successievelijk in de tijd. Bij N waarnemingslocaties in de ruimte zijn er even-

zovele tijdreeksen, en als er op k tijdstippen met interval Δt wordt waargenomen, dan zijn er eveneens k ruimtelijke velden. De aandacht richt zich bij monitoring op het detecteren van hoge waarnemingen en het volgen van trends. Omdat het aantal waarnemingen gewoonlijk vrij groot is (monitoring op 20 locaties met een interval van 4 weken levert jaarlijks liefst 260 waarnemingen) spelen statistische methoden een belangrijke rol. Deze kunnen gebruikt worden bij het detecteren van extreme waarden en bij het analyseren van korte en lange termijn trends. Maar ook is het opzetten van een monitoring schema gebaat bij statistische methoden. Met name gaat het daarbij om het vaststellen van het aantal waarnemingspunten in de ruimte en het vaststellen van het aantal monitoringdagen in de tijd. Tot op heden is het aantal beschreven monitoring studies in de bodemsanering gering. In andere gebieden, zoals in hydrologische studies naar de effecten van zure depositie op de samenstelling van de bodem zijn meer voorbeelden bekend.

Ruimtelijke variatie.

Een milieuparameter, zoals de concentratie aan cadmium of PAK's, varieert

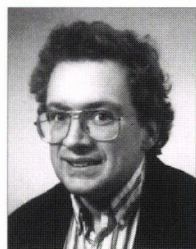
op ieder tijdstip in de ruimte. Immers, de waarnemingen zijn onderhevig aan toevallige fluctuaties, veroorzaakt door locale verschijnselen in de bodem, door toevalligheden bij het deponeren van de stof en door waterbeweging in de bodem. In het algemeen is het onmogelijk een volledig beeld te krijgen van deze variatie. Met een beperkt aantal metingen is het slechts mogelijk om de verontreinigingspatronen te inventariseren en te visualiseren. Daarom is er voor een kwantitatieve analyse van ruimtelijke variatie belangstelling ontstaan voor geostatistische methoden, waarbij de plaats van de waarnemingen een essentiële rol speelt.

Bij het beschrijven van ruimtelijke variabiliteit wordt veelal gebruik gemaakt van het semivariogram¹. Hierbij wordt een model gebruikt dat gekarakteriseerd wordt door vier essentiële kenmerken. Het eerste essentiële kenmerk heeft betrekking op *toevallige* fluctuaties, die b.v. veroorzaakt worden door de meetfout van het instrument. De variantie hiervan, het zg. *nugget effect*, speelt alleen een lokale rol: een meetfout op de ene locatie is niet gecorreleerd met de meetfout op een ander locatie. Daarnaast is er sprake van *ruimtelijke* correlatie: waarnemingen vlak bij elkaar lijken sterker op elkaar dan waarnemingen verder van elkaar af.

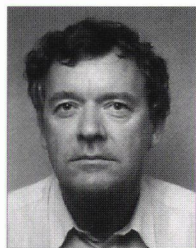
Deze afhankelijkheid vermindert met de afstand tussen de waarnemingslocaties volgens een model, bijvoorbeeld een exponentieel model. De vorm van het model is het tweede essentiële kenmerk. Op een afstand die gelijk is aan de range is de ruimtelijke correlatie volledig verdwenen, of nog maar 5% van de ruimtelijke afhankelijkheid op zeer korte afstand. Deze range is het derde essentiële kenmerk. De waarde die bij de range bereikt wordt is de variantie, ofwel de drempelwaarde en is het vierde essentiële kenmerk.

Er zijn tenminste 100 waarnemingen nodig om een betrouwbare schatting te krijgen van het semivariogram. Dit aantal kan als leidraad dienen bij het vaststellen van het aantal monitoring locaties. Als er meer waarnemingen beschikbaar komen in de loop der tijd en stationariteit in de tijd kan worden aangenomen

Over de auteurs:



Dr. ir. A. Stein,
is universitair hoofddocent in de geostatistiek aan de Landbouwniversiteit Wageningen.



Prof. dr. ir. J. Bouma,
is Hoogleraar in de bodeminventarisatie en de landevaluatie aan de Landbouwniversiteit Wageningen.

dan kan volstaan worden met minder monitoring locaties.

De achterliggende gedachte is dat de optredende verschillen van een bodem- of grondwatervariabele gebonden zijn aan specifieke locaties. De vraag of het gehalte aan een zwaar metaal ergens op een perceel een norm overschrijft is vaak even belangrijk als de vraag of de gemiddelde waarde deze norm overschrijft. De belangstelling gaat zowel uit naar de gemiddelde verontreiniging op een perceel, als naar verontreiniging op individuele punten.

Geostatistische methoden kunnen gebruikt worden voor het uitvoeren van

voorspellingen op niet bezochte locaties (kriging) en voor het bepalen van kansen op overschrijdingen van normen. Een eigenschap van kriging is dat 'smoothing' optreedt: een met kriging vervaardigde kaart toont minder variatie dan de werkelijkheid. Die variatie kan wel worden nagebootst m.b.v. stochastische simulaties: bij een gegeven (beperkt) aantal waarnemingen wordt nagegaan hoe volgens een toevalverdeling de ruimtelijke verdeling er uit kan zien. Dit is dan een afspiegeling van het toevalsveld zoals dat gedefinieerd is met behulp van gemiddelde, standaardafwijking en semivariogram.

Als waarnemingen beschikbaar zijn, dan kunnen simulaties worden uitgevoerd, geconditioneerd op de waarnemingen. Conditionele simulaties kunnen nuttig zijn bij het bepalen van waarnemingsschema's. In het eerste voorbeeld zal aandacht besteed worden aan stochastische simulaties en het effect daarop van het aantal waarnemingen.

Temporele variatie

In eerste instantie lijkt het goed denkbaar om waarnemingen in de tijd te behandelen als waarnemingen die in de ruimte langs een transect verzameld zijn en dan geostatistische methoden te gebruiken. Fluctuaties in de tijd verschillen echter essentieel van fluctuaties in de ruimte. Om te beginnen is er sprake van oorzaak en gevolg: een waarneming op tijdstip t heeft wel invloed op waarnemingen op tijdstip $t+1$, maar omgekeerd geldt dat niet. In de tweede plaats kennen waarnemingen in de tijd seizoensfluctuaties: het grondwater staat in de winter hoger dan in de zomer, met corresponderende opwaartse en neerwaartse fluxen; het nitrificerend vermogen van de bodem verandert als een functie van de temperatuur. Daarom moet er nadrukkelijk met seizoensfluctuaties worden rekening gehouden, zeker als monitoring systemen zich over verscheidene jaren uitstrekken.

Het noodzakelijke aantal waarnemingen in de tijd wordt gegeven door de zg. Nyquist frequentie: om een periodiciteit vast te kunnen stellen moet er tenminste twee keer per periode worden gemeten. Zo kan een 24-uurs fluctuatie nooit worden vastgesteld als er twee keer per jaar gemeten wordt. Omgekeerd, als de interesse bestaat uit het vastleggen van een seizoensfluctuatie dan is het zinloos om 2 keer per maand te gaan meten. Een meetfrequentie van eens in de maand kan

geschikt zijn om langjarige trends te detecteren. Als er voldoende waarnemingen zijn, dan kan een standaard tijdreeksanalyse worden uitgevoerd².

Variabiliteit in ruimte en tijd

Een gecombineerde analyse in ruimte en tijd kan uitgevoerd worden langs verschillende wegen:

- 1) De tijd wordt beschouwd als een extra dimensie van de ruimte. Problemen betreffen met name het zoeken naar een goede afstandsmaat in ruimte en tijd: een afstand van 1m in de ruimte is niet *a priori* gekoppeld aan een bepaalde tijdsmaat; deze maten kunnen zeer verschillend zijn. De doorlooptijd van nitraat in de bodem ligt in de orde van grootte van 1m/jaar, maar kan voor NH_4^+ veel korter zijn³.
- 2) Veranderingen in de tijd worden gemodelleerd als een tijdreeks op meer locaties. Deze methode werkt bevredigend met niet te veel waarnemingslocaties maar lange tijdreeksen, die weinig ontbrekende waarden kennen. Een bekend voorbeeld is een analyse van de windenergie van Ierland⁴ met uurgegevens over een periode van 18 jaar.
- 3) Bij veel waarnemingen op weinig tijdstippen is het mogelijk een verandering van ruimtelijke structuur te bepalen met behulp van een dynamisch semivariogram. Deze methode is met succes toegepast om de verspreiding van plantenepidemieën te analyseren⁵. Omdat op ieder tijdstip een semivariogram moet worden geschat, zijn betrekkelijk veel waarnemingen noodzakelijk.
- 4) Kalman filtering wordt gebruikt als er tevens sprake is van deterministische proces informatie en is ondermeer toegepast bij het analyseren van grondwaterstanden met het hydrologische model MODFLOW⁶.

Voorbeelden

Ruimtelijke variatie, conditionele simulaties en bemonsteringsdichtheid

Het simuleren van stochastische velden en het effect van het aantal waarnemingen is getest op gegevens op een diepte van 1.5-2m beneden maaiveld van de sanering van het Gerjo-terrein in Vlijmen⁷. De cyanide waarden (totale aantal 109) varieerden van 0.2 mg/kg tot 1050 mg/kg, met een gemiddeld gehalte op deze diepte van 92.1 mg/kg. De standaardafwijking was 184.2; de waarde van de variatie coëfficiënt (CV

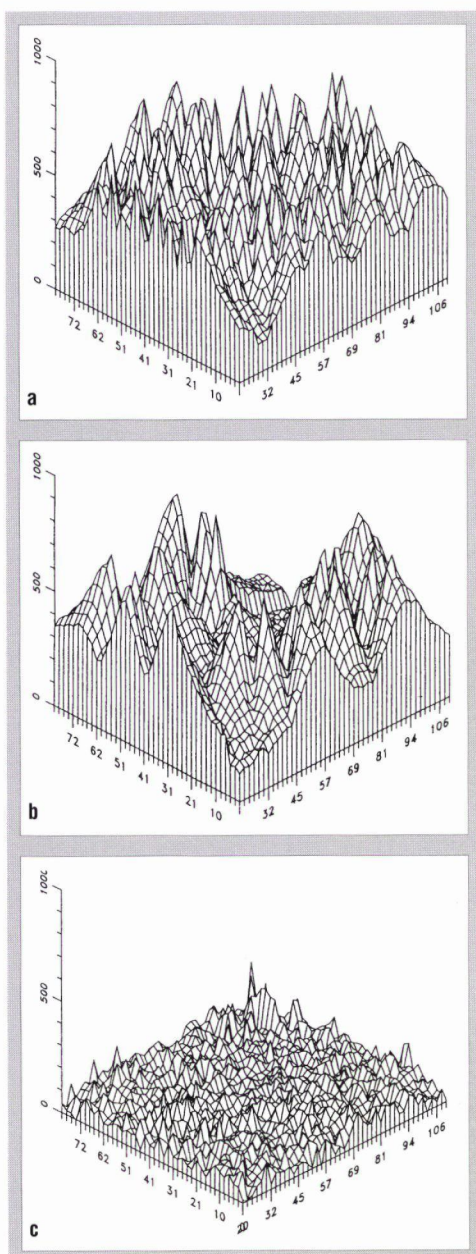


Fig. 1: Gegeneerde toevalsvelden met behulp van alle 109 waarnemingen (a), met behulp van 60 waarnemingen (b) en met behulp van 20 waarnemingen (c).

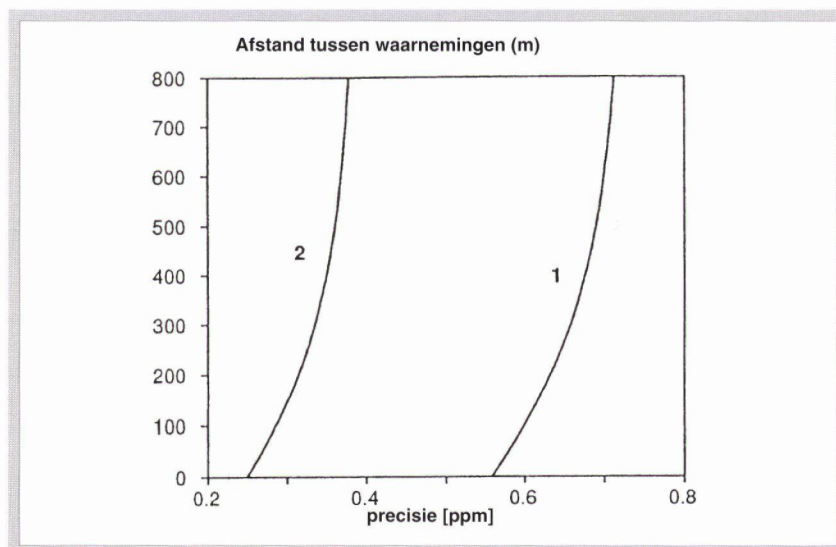


Fig. 2: Vereiste afstand tussen waarnemingen als een functie van het de gewenste precisie van het interpolatie oppervlak in Budel (1) en in Weert (2).

= 2) wijst op een scheve verdeling. Het semivariogram laat een exponentieel model zien. Het nugget effect was gelijk aan 25000, de sill-waarde aan 21700 en de range aan 19m. Van de ruimtelijke variatie bestaat een aanzienlijk gedeelte (ca. 50%) dus uit variantie op zeer korte afstand, veroorzaakt door binding van cyanide aan ijzer.

Van belang is het vaststellen van het aantal waarnemingen. Dat aantal hangt sterk af van het doel van de inventarisatie. Als er slechts één waarneming genomen wordt, dan kan deze van alles opleveren: bijvoorbeeld 0.2 mg/kg of 1050 mg/kg, of alles daartussen. Als er twee waarnemingen worden genomen, dan kan daarvan het gemiddelde genomen worden, maar de spreiding kan nog niet bepaald worden. Als er 20 waarnemingen genomen worden, dan kan met voldoende betrouwbaarheid de verdeling worden vastgesteld. Als er 100 waarnemingen genomen worden, dan kan het semivariogram bepaald worden. Met het toenemen van het aantal waarnemingen neemt de hoeveelheid verzamelde informatie ook toe. Om dat te visualiseren is in deze studie gekeken naar een conditionele simulatie waarbij alle 109 waarnemingen zijn gebruikt, een conditionele simulatie op basis van 60 waarnemingen en een conditionele simulatie op basis van 20 waarnemingen.

Als alle waarnemingen gebruikt worden (fig. 1a) dan treedt er variatie op gekoppeld aan de locatie, d.w.z. ruimtelijk van gedaante. Als er maar 60 waarnemingen worden gebruikt (fig. 1b), dan zien de velden er overeenkomstig uit, maar met minder ruimtelijke variatie.

Het patroon is rustiger, maar niet conform de realiteit, zoals blijkt in verge-

lijking tot fig. 1a. Het beeld wordt beoerd als het aantal punten wordt teruggebracht tot 20 (fig. 1c). De hoogte van het veld klopt niet meer (veroorzaakt door de loting van de 20 uit de 109 punten waarop geconditioneerd is), net zomin als de variatie binnen de velden. Op grond van 20 waarnemingen ontstaat een beeld dat weinig tot niets met de realiteit te maken heeft.

Ruimtelijke variatie, interpolatie en bemonsteringsdichtheid

In een gebied van 200 km² in de Kempen is Kriging gebruikt als een interpolatietechniek bij het onderzoek van cadmium verontreiniging⁸. Monsters zijn genomen in 2020 tuinen waarin ook het organische stof-gehalte en de pH zijn gemeten. Op kaarten is de kans aangegeven dat een kritisch cadmium gehalte van 2.5 mg Cd/kg wordt overschreden. Tevens is het minimum aantal bemonsteringspunten bepaald dat nodig

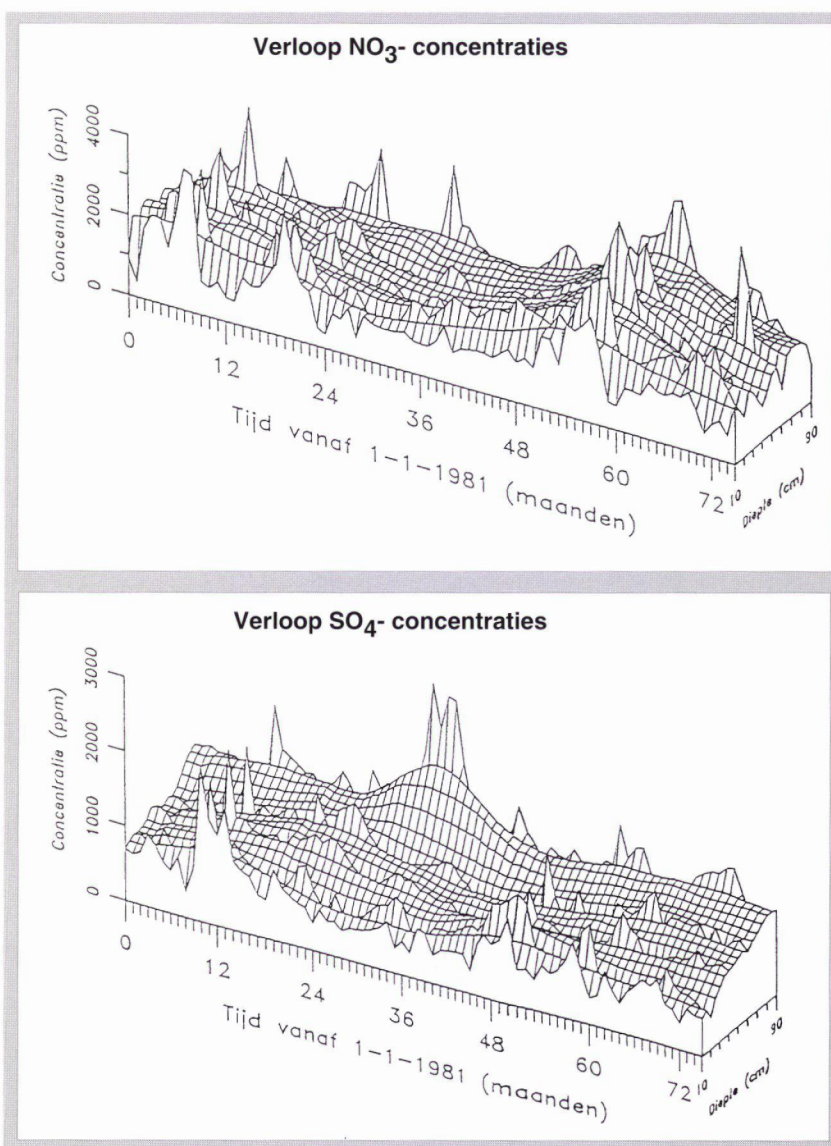


Fig. 3: Variatie in ruimte (diepte) en in tijd van NO₃- en SO₄²⁻ in de Hackfort locatie.

is voor het bereiken van een bepaalde nauwkeurigheid van voorspellingen.

Figuur 2 toont aan welke bemonsteringsdichtheid gewenst is om een geïnterpoleerde kaart van een bepaalde precisie te verwezenlijken, waarbij de twee gemeenten Budel en Weert afzonderlijk zijn bekeken. Om de huidige precisie te halen in Weert (1400 waarnemingen) kan met veel minder waarnemingen worden volstaan. Er zijn slechts voldoende waarnemingen nodig om het semivariogram te kunnen bepalen (ca. 100). Gevoegd bij de 600 waarnemingen in Budel betekent dit een reductie van het aantal waarnemingen met ca. 65%.

Temporele fluctuaties

In het studiegebied Hackfort in Oost-Gelderland heeft gedurende de periode 1981-1987 intensieve monitoring plaatsgevonden. Doel hiervan was om inzicht te krijgen in de veranderingen

van de bodem onder invloed van zure depositie⁹. Maandelijks zijn gegevens verzameld, o.a. aan nitraat, sulfaat, aluminium. Op 4 verschillende locaties met enigszins verschillende bodem- en vegetatiecondities is het bodemvocht verzameld en op zijn chemische samenstelling geanalyseerd. De dieptes 10cm, 20cm, 40cm, 60cm en 90cm beneden maaiveld zijn hierbij betrokken, vaak zijn metingen in duplo en (op 90 cm) in viervoud, verricht. De belangrijkste statistische vraag was of er sprake was van een significante trend in de concentraties in de loop der tijd.

Verschillen in temporele variabiliteit worden duidelijk uit de concentraties van NH_4^+ en SO_4^{2-} (fig. 3). Er is met name een duidelijke seizoensfluctuatie te zien met een periodiciteit van 12 maanden. De Nyquist frequentie leert ons dat 2 waarnemingen per jaar minimaal nodig zijn om zinvolle informatie te verkrijgen. Een dergelijke frequentie gaat voorbij aan de precisie die we zouden

willen hanteren om voorspellingen op niet waargenomen tijdstippen uit te voeren. Als deze precisie gegeven is, dan kan de waarnemingsdichtheid op grond daarvan worden vastgesteld¹⁰.

Optimalisatie van een bestaand freatisch meetnet in twee waterschappen

Een ruimtelijke analyse van de freatische stijghoogtes in de waterschappen Salland en Bezuiden de Vecht is een voorbeeld van meetnetoptimalisatie die op dit moment door de waterschappen wordt gehanteerd. Gegevens zijn beschikbaar van 4 verschillende dagen: 3 maart 1988, 3 juli 1988, 3 maart 1992 en 3 juli 1992, geselecteerd op basis van de hydrologische condities (1988 was een nat jaar, 1992 een droog jaar, in maart is er sprake van relatief hoge grondwater standen, in juli van lage). Het basisbestand betreft de peilbuizen waar tenminste op één van de genoemde dagen een waarneming is ver-

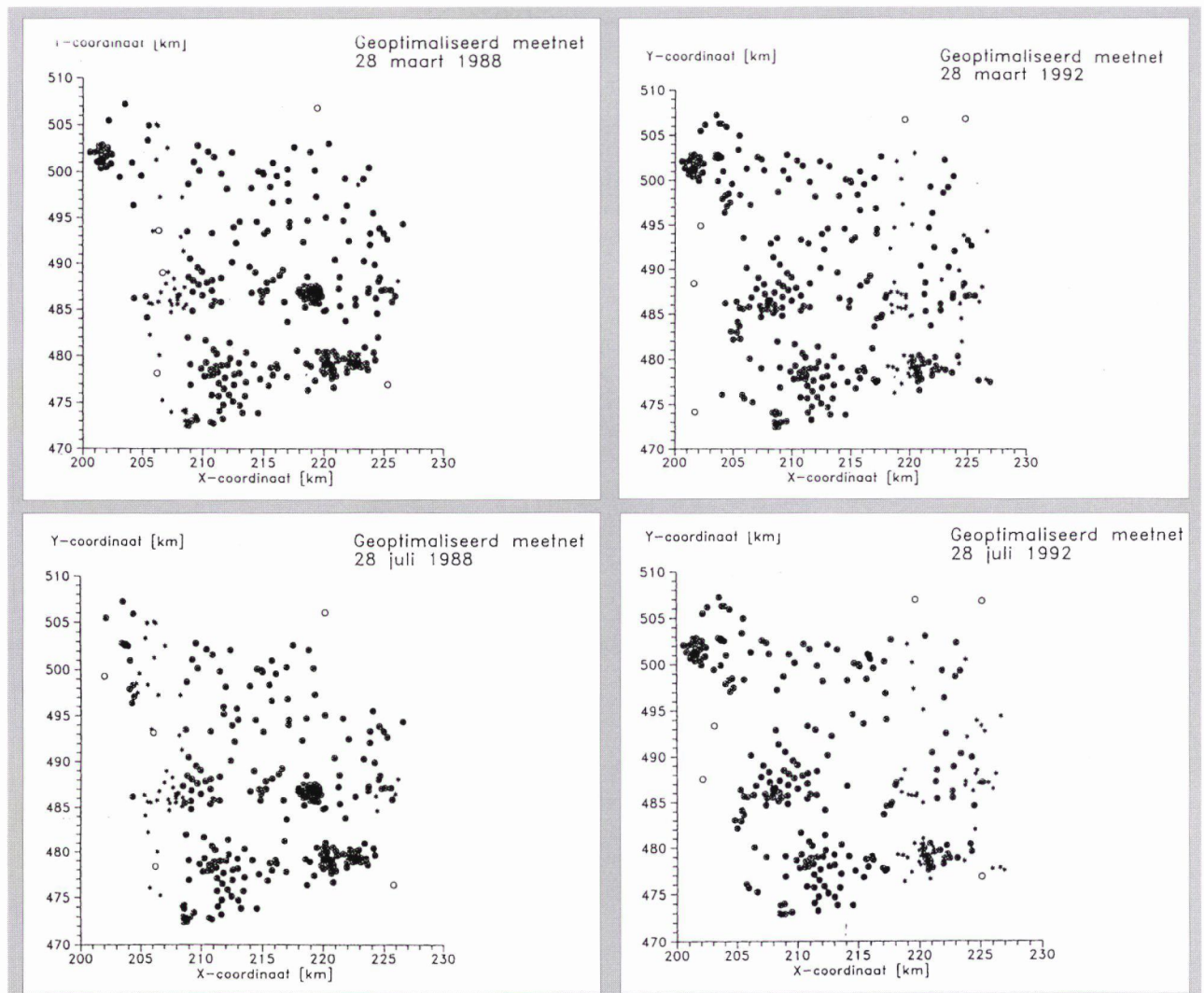


Fig. 5: Geoptimaliseerde meetnetten voor de verschillende monitoring dagen (●). De 50 peilbuizen die onvoldoende informatie opleveren zijn verwijderd (*) en 5 nieuwe peillocaties zijn toegevoegd (○)

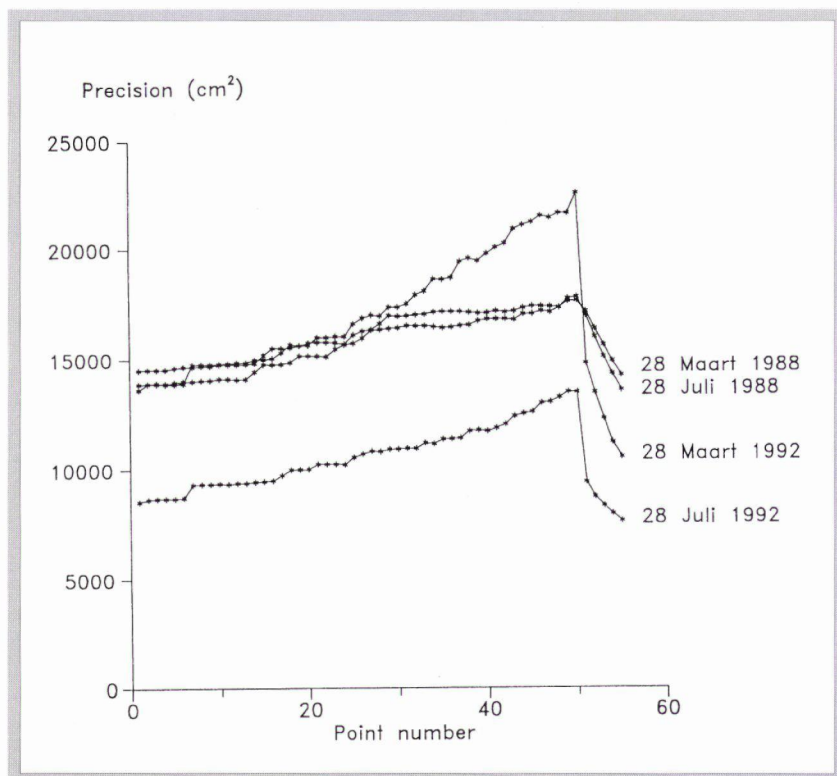


Fig. 4: Toe- en afname van de ruimtelijke onzekerheden bij het uitdunnen (stijgende lijn) en het verdichten van het meetnet. Duidelijk herkenbaar is dat de onzekerheid daalt onder de onzekerheid van het oorspronkelijke meetnet bij toevoeging van een vijfde peilbuis.

richt; dit bestand bestaat uit 415 peilbuizen in de twee waterschappen.

Optimalisatie bestaat zowel uit het verwijderen van peilbuizen (verdunning) die *in ruimtelijke zin* weinig bijdragen aan het verkrijgen van een kaart met uniforme betrouwbaarheid, als uit het toevoegen van peilbuizen op die locaties waarmee een betere betrouwbaarheid verkregen kan worden. Zowel het effect van verdunning als van uitbreiding van het huidige meetnet is geanalyseerd door gebruik te maken van de gemiddelde kriging variantie in het studiegebied. De resultaten staan weergegeven in figuur 4. Het elimineren van 50 peilbuizen en het toevoegen van 5 nieuwe peilbuizen, leidt telkens tot een *toename van de precisie*. Deze toename is soms gering (enkele procenten op beide dagen in 1988) tot zeer fors (23 procent op 28 juli 1992). Met de gehanteerde aannames leidt dit optimalisatieproces op een totaal van 415 peilbuizen al tot een kostenbesparing van 11%.

In figuur 5 zijn de configuraties van de optimale meetnetten gegeven, waarbij de vijf nieuwe meetlocaties zijn ingetekend. Uitdunning op de verschillende dagen vindt plaats op verschillende punten. In 1988 liggen de punten die de kleinste bijdrage geven aan de

precisie van de voorspellingen aan de westkant van het studiegebied. In 1992 zijn het veeleer de peilbuizen in het oosten van het studiegebied die verwijderd kunnen worden. Vanuit een praktische invalshoek treedt een complicatie op: een optimaal meetnet voor een *natte* periode is anders dan een optimaal meetnet voor een *droge* periode.

Deze analyse is een eerste stap om te komen tot een daadwerkelijke meetnetoptimalisatie. Nog niet alle informatie is hierbij gebruikt, en er zijn nog enkele haken en ogen aan de gevolgde procedure. In principe lijkt de geschetste procedure aantrekkelijk, omdat er nu al sprake is van een aanzienlijke reductie in opname capaciteit. Additionele randvoorwaarden voor volledige optima-



lisatie zijn: een aantal punten moet bevestigd blijven; het bepalen van de ruimtelijke onzekerheid van het netwerk dient niet zozeer op basis van toevallige punten te gebeuren, maar op basis van een aantal van tevoren vastgestelde punten; de kosten van monitoring laten een bepaald maximum aantal punten toe, terwijl er mogelijk beperkingen zijn op een mogelijke localisatie van meetpunten (die nu nog op basis van toeval geselecteerd zijn), etc. Ook kan beschikbare voorinformatie beter gebruikt worden.

Literatuur

- 1) **N.A.C. Cressie, 1991**
Statistics for spatial data. Wiley, New York.
- 2) **C. Chatfield, 1989**
The analysis of time: an introduction - 4th ed.
- 3) **N. van Breemen, W.F.J. Visser en Th. Pape, 1988**
Biogeochemistry of an oak-woodland ecosystem in the Netherlands affected by acid atmospheric deposition. Agric. Research Reports 930, Pudoc, Wageningen, 198pp.
- 4) **J. Haslett en A.E. Raftery, 1989**
Space-time modelling with long-memory dependence: assessing Ireland's wind power resource (with discussion). Applied Statistics 38, 1-50.
- 5) **A. Stein, C. Kocks, J.C. Zadoks, H. Frinking, T. Ruissen & D.E. Meyers**
A geostatistical analysis of the spatio-temporal development of downy mildew epidemics in cabbage. Aangeboden aan Fytopathology.
- 6) **F.C. van Geer, C.B.M. te Stoet en Zhou Yangxiaoyao, 1991**
Using Kalman filtering to improve and quantify the uncertainty of numerical groundwater simulations. 1. The role of system noise and its calibration. Water Resources Research 27, 1987-1994.
- 7) **I.G. Staritsky, P.H.M. Sloot en A. Stein, 1992**
Spatial variability and sampling of cyanide polluted soil on former galvanic factory premises. Water, Air, and Soil Pollution 61, 1-16.
- 8) **I.G. Staritsky en A. Stein, 1990**
Cadmium verontreiniging in Budel en Weert: een geostatistische studie. Landbouwwerksite Wageningen, intern rapport.
- 9) **A. Stein en N. van Breemen. In druk**
Time series analysis of changes in the soil solution: evidence for approaching nitrogen saturation in Dutch forest soils. Agriculture, Ecosystems and the Environment.
- 10) **P.J.W.M. Miiskens, 1978**
The use of autocorrelation techniques for selecting optimal sampling frequency - application to surveillance of surface water quality. Analytica Chimica Acta 103, 445-457.