

Statistiek bij regionale proevenverzamelingen

Regionale proevenverzamelingen lenen zich bij uitstek om de ruimtelijke structuur van variaties van een grondeigenschap binnen een uitgestrekte grondlaag te analyseren. Een veel voorkomende toepassing is het schatten van de karakteristieke waarde van lokale laaggemiddelden van een grondeigenschap voor grondmechanische ontwerpanalyses voor lijninfrastructuurprojecten. In een voorgaand artikel [1] is, aan de hand van een praktijkvoorbeeld, een ruimtelijk variatie model afgeleid voor de statistische beschrijving van een proevenverzameling. In dit artikel gaan we nader in op de toepassing.

Resumé van het ruimtelijk statistisch model

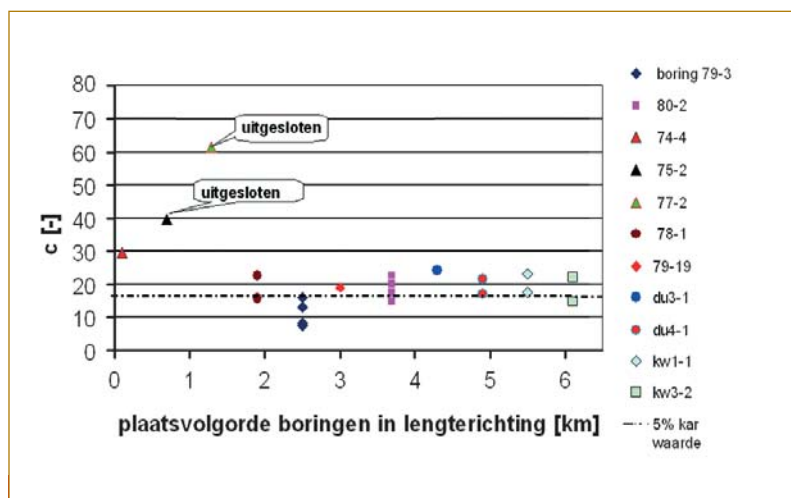
De proevenverzameling, aan de hand waarvan een ruimtelijk statistisch model werd geconstrueerd, is weergegeven in *figuur 1*. Het gaat om waarnemingen van de samendrukkingsparameter C , bepaald uit laboratoriumproeven op grondmonsters die zijn ingewonnen uit 11 boringen langs een dijk, over een afstand van ongeveer 6 kilometer. Bij de analyse in het voorgaande

artikel zijn voornamelijk hoge 'uitbijters' in de reeks waarnemingen (in boring 75-2 en boring 77-2) buiten beschouwing gelaten. Verderop in dit artikel zullen we het effect van het al dan niet meenemen nader onderzoeken.

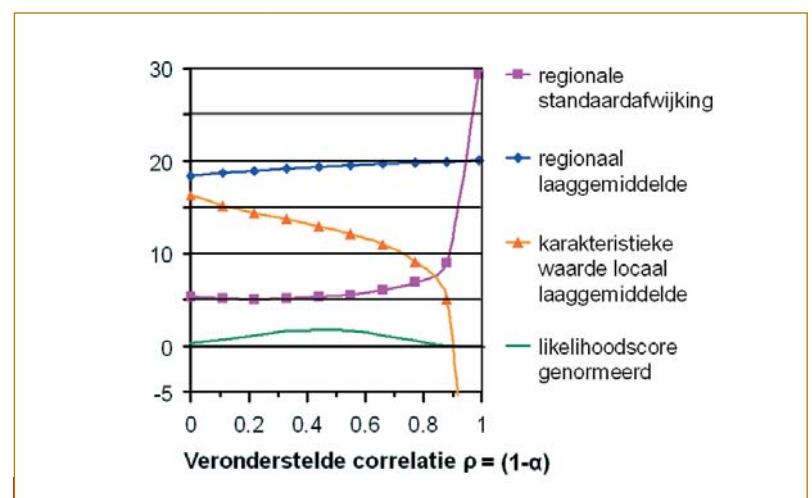
In het model onderscheiden we het 'locale laaggemiddelde', de 'locale variantie', het 'regionale laaggemiddelde' en de 'regionale variantie'. Het locale laaggemiddelde is het laaggemiddelde van de grondeigenschap op een locatie. Om de gedachten te bepalen: het gemiddelde van de grondeigenschap in één boring in *figuur 1* is een lokaal laaggemiddelde. Dit laaggemiddelde is representatief voor een gebied rondom de boring van enige tientallen meters. De locale variantie is de variantie van afwijkingen van de 'puntwaarden' van de grondeigenschap op een locatie ten opzichte van het locale laaggemiddelde. Het regionale laaggemiddelde is het laaggemiddelde over de gehele grondlaag (dus over alle boringen in *figuur 1*). En de regionale variantie is de variantie van afwijkingen van puntwaarden in het gehele gebied, ten opzichte van het regionale

laaggemiddelde. 'Puntwaarden' zijn in dit verband waarden die gemeten worden aan kleine grondvolumes, met behulp van in situ terreinproeven of in het laboratorium. De verhouding α tussen locale variantie en regionale variantie is een belangrijke parameter van het ruimtelijke model. Deze verhouding kenmerkt de ruimtelijke correlatiestructuur van de proevenverzameling. In [1] is afgeleid dat bij dit model binnen een boring de waarnemingen van de grondeigenschap gecorreleerd zijn met correlatiecoëfficiënt $\rho = (1 - \alpha)$, terwijl waarnemingen in verschillende boringen (die ver genoeg uit elkaar liggen) onderling ongecorreleerd zijn. Dit leidt tot formules voor de schatting van regionale verwachtingswaarde, regionale standaardafwijking en karakteristieke waarde van lokale laaggemid-

de. 'Puntwaarden' zijn in dit verband waarden die gemeten worden aan kleine grondvolumes, met behulp van in situ terreinproeven of in het laboratorium. De verhouding α tussen locale variantie en regionale variantie is een belangrijke parameter van het ruimtelijke model. Deze verhouding kenmerkt de ruimtelijke correlatiestructuur van de proevenverzameling. In [1] is afgeleid dat bij dit model binnen een boring de waarnemingen van de grondeigenschap gecorreleerd zijn met correlatiecoëfficiënt $\rho = (1 - \alpha)$, terwijl waarnemingen in verschillende boringen (die ver genoeg uit elkaar liggen) onderling ongecorreleerd zijn. Dit leidt tot formules voor de schatting van regionale verwachtingswaarde, regionale standaardafwijking en karakteristieke waarde van lokale laaggemid-



Figuur 1 Grafische representatie van de waarnemingen; de boringen zijn van links naar rechts in de juiste volgorde binnen de strekking uitgezet. De streepstippelijne geeft de karakteristieke waarde van het gemiddelde van de grondlaag, volgens de 'klassieke' formules in [1].



Figuur 2 Berekening van regionaal gemiddelde, regionale standaardafwijking en karakteristieke waarde voor lokale laaggemiddelden van C , gebaseerd op de proevenverzameling in *figuur 1*, als functie van de correlatie tussen waarden van C binnen een boring $\rho = (1 - \alpha)$. De waarden voor $\rho = 0$ geven berekeningsresultaten met de klassieke statistische formules voor ongecorreleerde waarnemingen weer.

delden die afwijken van de doorgaans gebruikte klassieke statistische formules voor ongecorrigeerde waarnemingen. Resultaten van berekeningen met deze nieuwe formules, op basis van de waarnemingen in *figuur 1*, zijn weergegeven in *figuur 2*. Het regionale gemiddelde, de regionale standaardafwijking en de karakteristieke waarde van het locale laaggemiddelde zijn functies van de aangenomen waarde voor de variantieverhouding α en dus de correlatie binnen een boring, $\rho=(1-\alpha)$. Voor een nadere uitleg en de wiskundige onderbouwing wordt verwezen naar [1].

Schatting van α ; een Bayesiaanse benadering

De grootte van de variantieverhouding α is in beginsel voor elke proevenverzameling verschillend en moet aan de hand van de waarnemingen zelf geschat worden. In *figuur 2* is het verloop van de *likelihood*-functie weergegeven. Dit is een functie die de waarde de gezamenlijke kansdichtheid van de waarnemingen weergeeft, uitgaande van een aangenomen waarde van de correlatie $\rho=(1-\alpha)$ van waarnemingen binnen een boring. We interpreteren deze functie als volgt: de *likelihood*-functie geeft aan wat de waarschijnlijkheid is van de waarde van de correlatie $(1-\alpha)$ op basis van de waarnemingen zelf. Dit noemen we een Bayesiaanse interpretatie. De 'genormaliseerde' *likelihood*-functie noemen we de α posteriori kansdichtheidsfunctie van $(1-\alpha)$. Voor de geïnteresseerde lezer is de wiskundige achtergrond en verdere uitwerking weergegeven in het kader 'Bayesiaanse benadering' (zie ook [2]).

Aan de *likelihood*-functie in *figuur 2* zien we dat correlaties van 0,3 tot 0,6 de hoogste scores hebben. Dat betekent dat op grond van de waarnemingen een waarde van ρ van 0,4 tot 0,7 veel waarschijnlijker is dan de waarde $\alpha=1$ (corresponderend met $\rho=0$), waarbij de klassieke statistische formules (zie [1]) gelden om verwachtingswaarde, standaardafwijking en karakteristieke waarde van locale gemiddelden te schatten.

De *likelihood*-functie is echter zodanig dat er niet echt één specifieke waarde voor α uitspringt. Dit komt onder andere doordat het aantal waarnemingen (20 stuks) niet erg groot is, waardoor het onderscheidend vermogen van een statistische berekeningsmethode beperkt is. Ook zal dit het gevolg zijn van het feit dat de fysische werkelijkheid grilliger is dan het ideaalmodel dat we voor de beschrijving ervan hanteren. Om die reden, en omdat dit vanuit een Bayesiaanse optiek ook logischer is, nemen we de hele α posteriori kansverdeling mee in de bepaling van de karakteristieke waarde van locale laaggemiddelden. Dit doen we via integratie van het product van de karakteristieke waarde bij een bepaalde $(1-\alpha)$ en de hiermee corresponderende α posteriori kansdichtheid. In het kader is de wiskundige uitwerking gegeven.

Berekeningsresultaten, effect van meenemen van de uitbijters

De uitkomst van deze integratie levert een karakteristieke waarde voor locale gemiddelden van 12,5. De schatting van de karakteristieke waarde, berekend volgens de 'klassieke'

berekening was 16,4 (zie *figuur 1*). We herinneren er aan dat bij deze berekeningen de 'hoge' uitbijters (boringen 75-2 en 77-2) zijn weggelaten. Wanneer we die wel in de analyse meenemen blijkt dat ze een grote invloed op de berekende karakteristieke waarde hebben. In *tabel 1* zijn de berekende karakteristieke waarden van locale gemiddelden weergegeven voor het geval de hoge uitbijters wel worden meegenomen. Daarbij is uit gegaan van de aanname dat de grondeigenschap normaal verdeeld is.

Tabel 1: Berekende karakteristieke waarden voor locale laaggemiddelden met al dan niet meenemen van uitbijters, op basis van aanname van een normale verdeling

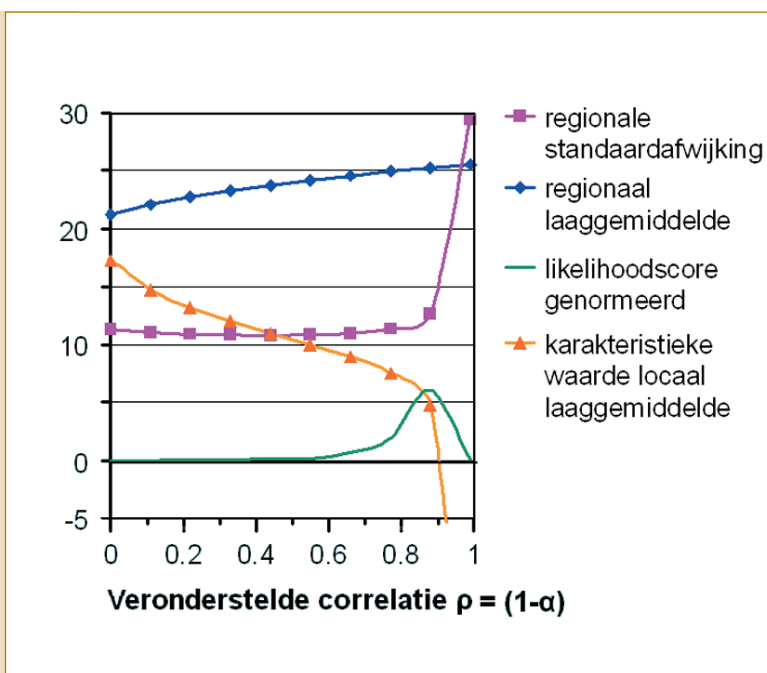
Selectie van waarnemingen

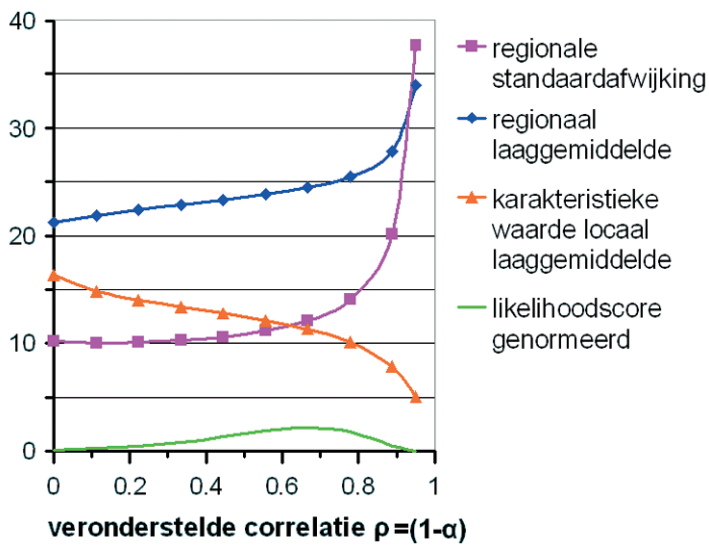
	Karakteristiek locale laaggemiddelde
Alle, uitgezonderd boring 75-2 en 77-2	12,8
Alle, uitgezonderd boring 77-2	11,0
Alle	5,9

Het is opvallend dat het toevoegen van 'gunstiger' waarnemingen tot een ongunstiger schatting van de karakteristieke waarde leidt. Vooral het effect van de extreem hoge waarneming in boring 77-2 is desastreus voor de schatting van de karakteristieke waarde van het locale laaggemiddelde. Daar zijn verschillende oorzaken voor. Ten eerste wordt door het meenemen van uitbijters de regionale variantie groter. Maar ook neemt hierdoor de verhouding α tussen de locale en de regionale variantie af en neemt dus de correlatie $\rho=(1-\alpha)$ binnen de boringen toe. In *figuur 3* zien we dat aan een verschuiving naar rechts van de grote scores van de *likelihood*-functie, ten opzichte van *figuur 2*. Daardoor krijgen de ongunstige waarden van de karakteristieke locale laaggemiddelden een grote invloed in de integratie.

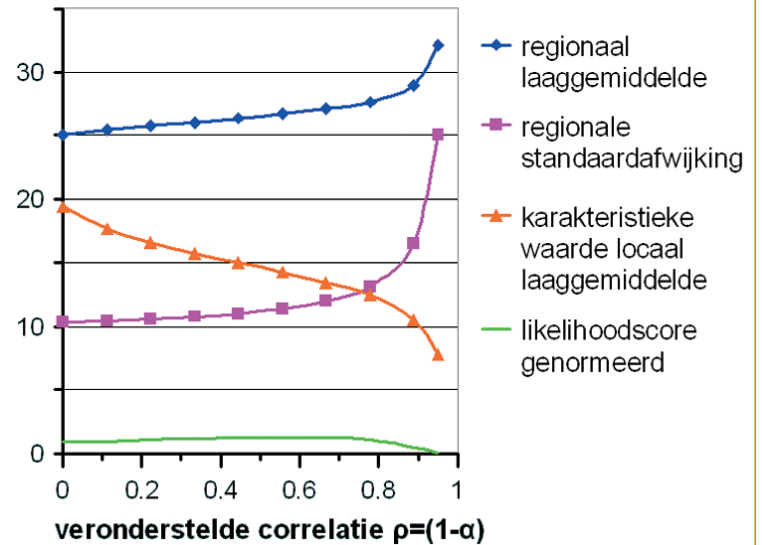
Zoals in het voorgaande artikel [1] al is gesteld moeten uitbijters altijd nader onderzocht worden. Als niet aannemelijk te maken is waarom een uitbijter niet tot de te karakteriseren populatie behoort, is er ook geen reden om die uit te sluiten bij de statistische analyse van de proevenverzameling. Veronderstel dat dit laatste in ons voorbeeld het geval is. In *figuur 3* zien we dat het meenemen van de uitbijters, in combinatie met de aanname van een normale verdeling, bij hoge correlaties tot negatieve en dus fysisch onmogelijke karakteristieke locale gemiddelden leidt. Omdat bij deze correlaties de *likelihood*-scores

Figuur 3 Als *figuur 2*, maar dan met meenemen van de hoge uitbijters in de reeks waarnemingen. Opmerkelijk is dat de top van de *likelihood*-functie naar rechts is verschoven. Dit komt omdat het meenemen van de hoge uitbijters ervoor zorgt dat de verhouding α van locale en regionale variantie afneemt. Immers, de variatie binnen een boring wijzigt niet, terwijl de regionale variatie wel toeneemt.





Figuur 4 Als figuur 3, maar dan met aanname van de lognormale kansverdeling voor de grondeigenschap C



Figuur 5 Als figuur 4, maar met weglaten van de waarnemingen in boring 79-3 en 80-2

groot zijn, hebben deze negatieve waarden wel een grote invloed op de 'uitgeïntegreerde' karakteristieke waarde (berekend volgens vergelijking 1.2 in het kader 'Bayesiaanse benadering'). De statistische analyse wordt daardoor, in tegenstelling tot in figuur 2, sterk beïnvloed door fysisch inconsistente uitkomsten. Om die reden wordt nu de aanname van een normale verdeling voor de waarnemingen verworpen

Aanname lognormale kansverdeling voor waarnemingen

Wanneer we uit gaan van een lognormale verdeling wordt in elk geval vermeden dat fysisch onmogelijke karakteristieke waarden worden berekend. In het voorgaande artikel [1] is al aangetoond dat de lognormale verdeling, net als de normale verdeling, vanuit statistisch oogpunt niet kan worden verworpen. We gaan daarom nu na wat het resultaat van de statistische analyse wordt als we een lognormale verdeling voor de waarnemingen aannemen.

De analyse onder aanname van een lognormale verdeling voor de waarnemingen verloopt geheel analoog aan de analyse bij aanname van een normale verdeling. Alleen worden de statistische bewerkingen niet uitgevoerd op de waarnemingen zelf, maar op de logaritmen van de waarnemingen. We vinden dan de karakteristieke waarde van lokale laaggemiddelden van de logaritme van de grondparameter. Door vervolgens hiervan de exponent te nemen vinden we

de gezochte karakteristieke waarde. In tabel 2 zijn de berekende karakteristieke waarden voor lokale laaggemiddelden weergegeven voor zowel de aanname van normale als lognormale verdeling. Het effect, dat hoge uitbijters leiden tot lagere karakteristieke waarden is er ook bij aanname van de lognormale verdeling, maar veel minder desastreus.

Tabel 2: Berekende karakteristieke waarden voor lokale laaggemiddelden, al dan niet met meenemen van hoge uitbijters, op basis van aanname van de normale en de lognormale verdeling

Selectie van waarnemingen	Aangenomen kansverdelingstype	
	Normaal	Lognormaal
Alle, uitgezonderd		
75-2 en 77-2	12,8	12,4
Alle, uitgezonderd 77-2	11,0	12,3
Alle	5,9	11,6

In figuur 4 is het verloop van de schatters voor regionaal gemiddelde, de regionale standaardafwijking en de karakteristieke ondergrens als functie van de aangenomen correlatie binnen een boring. Bij deze berekening zijn alle waarnemingen, dus ook de uitbijters, meegenomen.

Deze analyse rechtvaardigt de conclusie dat voor grondeigenschappen die per definitie een posi-

tieve waarde hebben de aanname van een lognormale kansverdeling de voorkeur geniet boven de aanname van een normale verdeling. Het blijft natuurlijk zaak om zo'n aanname te toetsen op aannemelijkheid, maar dit geldt eigenlijk ook voor de aanname van normaliteit. In de praktijk wordt er doorgaans van uitgegaan dat voor aanname van normaliteit die toets niet nodig is. Uit bovenstaande analyse blijkt dat dit een onterecht gebruik is.

In tabel II zien we dat door aanname van de lognormale verdeling de invloed van enkele hoge uitbijters op de berekende karakteristieke (5%-ondergrens) waarden van lokale laaggemiddelden nogal beperkt is. Door de oogharen heen kijkend zou je dit ook verwachten. Hoge uitbijters geven immers weinig informatie over statistische ondergrenzen van een aan de onderkant fysisch begrensde populatie. Het zou dus ook onlogisch zijn om te verwachten dat het meenemen van een beperkt aantal hoge uitbijters zou leiden tot substantieel lagere (of hogere) ondergrenswaarden.

Opzet van onderzoek voor een regionale proevenverzameling

De regionale proevenverzameling van figuur 1 heeft ons op het spoor gezet van het ruimtelijke variatiemodel, waarbij de spreiding binnen een locatie, ten opzichte van het lokale laaggemiddelde, kleiner is dan de totale spreiding in de hele regionale verzameling. Dat kwam omdat er

een aantal boringen is waarin meerdere waarnemingen aanwezig zijn, die in feite de informatie moeten verschaffen over de verhouding tussen locale en regionale variatie van de waarnemingen. Zou bij geen van de boringen meer dan één waarnemingen aanwezig zijn, dan bevat de proevenverzameling uiteraard ook geen informatie over locale variatie. Hebben we slechts enkele boringen met twee waarnemingen, dan bevat de verzameling wel wat informatie over locale spreiding, maar die wordt dan erg vaag. We kunnen dit demonstreren door in de statistische bewerking van de proevenverzameling de boringen 79-3 en 80-2 weg te laten. Dit zijn de boringen met een wat groter aantal waarnemingen, die derhalve de meeste informatie over de locale variatie bevatten. *Figuur 5* geeft de berekeningsresultaten weer zoals in *figuur 4*, maar dan met weglaten van deze twee boringen. We zien dat als gevolg hiervan de *likelihood*-scores meer uniform verdeeld zijn in de correlatierange van 0 - 0,8. Dat wil zeggen dat waarnemingen weinig informatie bevatten over de variantieverhouding α . De bijbehorende schatting van de karakteristieke waarde is nu 14,76. Deze schatting is gunstiger dan de schatting waarbij alle waarnemingen zijn meegenomen (namelijk 11,5, zie tabel 2), en laat dus zien dat het niet inwinnen van informatie over de verhouding van locale en regionale kan leiden tot onveilige schattingen van de karakteristieke waarde van locale laaggemiddelden.

Een belangrijke conclusie die we hier uit trekken is dat bij het opzetten van grondonderzoek, om een regionale proevenverzameling op te stellen, er altijd voor gezorgd moet worden dat er voldoende waarnemingen zijn om de verhouding tussen locale en regionale variantie, te bepalen. Een proevenverzameling waarin boringen met meerdere grondmonsters per boring ontbreken is een onbetrouwbare proevenverzameling om karakteristieke schattingen van locale laaggemiddelden te bepalen. In beginsel is dus een opzet voor het grondonderzoek met relatief weinig boringen, maar met relatief veel, bijvoorbeeld vier tot zes, waarnemingen per boring veruit te prefereren boven een opzet met veel boringen, maar met slechts één of hooguit twee waarnemingen per boring.

Regionale proevenverzameling volgens de TAW-Leiraad Rivierdijken

Het gebruik van regionale proevenverzamelingen voor de schuifsterkte in grondlagen is in de Leidraad voor het Ontwerpen van Rivierdijken (deel 2) geïntroduceerd [3]. Bij het berekenen van statistische kenmerken van de proevenverza-

melingen wordt geen ruimtelijk correlatiemodel gebruikt. Het berekenen van regionale gemiddelden en varianties van de proevenverzameling gebeurt met de klassieke formules. Het ruimtelijke model wordt echter wel gebruikt om karakteristieke waarden voor locale gemiddelden te berekenen. Daarbij wordt de volgende formule gebruikt:

$$C_{loc\ gem, kar} = \hat{\mu}_C - t_{N_{tot}-1}^{0,95} \hat{\sigma}_{C, reg} \sqrt{\frac{1}{N_{tot}} + I^2} \quad (1)$$

Hierin is I^2 de factor waarmee het effect van spreiding van locale gemiddelden in rekening gebracht wordt. Hiervoor wordt in de TAW-leidraad $I^2 = 0,25$ aangehouden. Deze keuze komt overeen met $\alpha = 0,75$ (vergelijk formule 1 met formule 1.11 van bijlage 1 van het voorgaande artikel [1]). Dit betekent dat in de TAW-leidraad (impliciet) uitgegaan is van het idee dat binnen één locatie de variantie van schuifsterkte van proefmonsters zo'n 75% van de regionale variantie in de proevenverzameling bedraagt. Kijken we naar de figuren 2 t/m 5, dan zien we dat α -waarden waarden van 0,2-0,5 voor de in *figuur 1* beschouwde proevenverzameling waarschijnlijk zijn.

De keuze in de Leidraad, zo'n twintig jaar geleden, was een educated guess, die niet gebaseerd is op een statistische analyse. Mogelijk zou een ruimtelijke analyse, zoals hierboven beschreven, van de destijds beschikbare proevenverzamelingen van schuifsterkteparameters tot andere keuzen hebben geleid; hierover kunnen we vooralsnog slechts speculeren. Het lijkt wel zinvol dit alsnog te doen.

Gebruiken we formule (1) (met $t = 1,65$) voor de proevenverzameling in *figuur 1* dan vinden we karakteristieke waarden voor locale gemiddelden van 11,2 (waarbij alle waarnemingen zijn meegenomen). We zien dat de berekende karakteristieke waarde van locale laaggemiddelden niet eens zo veel afwijkt van die in *tabel 2* (als daarbij uitgegaan wordt van een lognormale kansverdeling). Omdat de onderliggende modellen nogal verschillen houdt de vergelijking hiermee op. Verdergaande conclusies zijn speculatief.

De formule (1) uit de Leidraad voor Rivierdijken is, als best practice, in de loop der jaren ook wel toegepast op regionale proevenverzamelingen van grondeigenschappen bij grote infrastructuurprojecten. De analyse in bovenstaande alinea's laat zien dat dit praktisch misschien zo gek nog niet was, maar roept tegelijkertijd de

vraag op of dit vanuit statistisch oogpunt wel correct was. Aanbevolen wordt om in nieuw voorkomende gevallen gebruik te maken van het in dit en het voorgaande artikel voorgestelde statistische model.

Karakteristieke waarden voor locale laaggemiddelden in tabel 1 van NEN6740

In tabel 1 van de geotechnische norm NEN 6740 worden indicaties voor aan te houden karakteristieke (onder- en bovengrens-) waarden van laaggemiddelden op een locatie aangegeven. Deze indicaties zijn ontstaan op basis van deskundigenoordeel bij het opstellen van de norm. Hierbij zijn weliswaar niet expliciet proevenverzamelingen gebruikt, maar de ervaring van de bij het opstellen van de norm betrokken deskundigen kan niettemin opgevat worden als (langjarige) cumulatie van praktijkervaringen en in die zin, als het ware, gebaseerd op landdekkende regionale proevenverzamelingen.

In de norm worden ook variatiecoëfficiënten van de in *tabel 1* benoemde grondeigenschappen genoemd, overigens zonder dat precies is aangegeven waar deze spreidingen precies betrekking op hebben. Verschillende interpretaties zijn mogelijk. Door de CUR-commissie C135 is een nadere duiding van *tabel 1* en de interpretatie van de bijbehorende variatiecoëfficiënten, en de hieruit voortvloeiende gebruiksmogelijkheid van *tabel 1*, bestudeerd [4]. In een afzonderlijk artikel in dit blad wordt hier nader op ingegaan [5].

Conclusies en aanbevelingen

Aan de hand van een regionale proevenverzameling is een ruimtelijk stochastisch model voor de karakterisering van ruimtelijke variatie van grondeigenschappen beschreven. Daarbij zijn statistische formules afgeleid, voor het schatten van karakteristieke waarden van locale laaggemiddelden van de grondeigenschap op een willekeurige locatie, rekening houdend met de hierin aanwezige ruimtelijke correlatiestructuur. De hiermee berekende karakteristieke waarde van locale laaggemiddelden sluit, in tegenstelling tot berekende karakteristieke waarden met de klassieke formules, goed aan bij de waarnemingen van de regionale proevenverzameling in *figuur 1*. Dit ruimtelijke model is beschreven in [1] en is ingegeven door stochastische modellering die gebruikelijk wordt toegepast in de geostatistiek.

Bij een stochastisch model moet een aanname worden gedaan over het type kansverdeling van de populatie en dus ook van waarnemingen in de proevenverzameling. Bij modellering van grond-

Een Bayesiaanse benadering

We willen de correlatie $\rho (= 1 - \alpha)$ schatten aan de hand van de waarnemingen zelf. Daartoe veronderstellen we een a priori kansdichtheidsverdeling voor ρ namelijk $f'_\rho(\xi) (0 \leq \xi \leq 1)$, die we straks nader zullen specificeren.

Voor het gemak noteren we de waarnemingen met de vector \hat{c} (zie ook kader in [1]) en we noteren de samengestelde kansdichtheid van de waarnemingen, gegeven een correlatie, als $f_c(\hat{c}; \xi)$. Volgens het theorema van Bayes kunnen we de zogenaamde à posteriori kansdichtheidsverdeling van ρ berekenen, gegeven de waarnemingen, met:

$$f''_\rho(\xi | \hat{c}) = f_c(\hat{c}; \xi) f'_\rho(\xi) J \quad (0 < \xi \leq 1) \quad (1.1)$$

Hierin is J een normaliserende constante, die

ervoor moet zorgen dat het linkerlid van vergelijking (1.1) ook een echte kansdichtheidsfunctie is, dat wil zeggen met een geïntegreerd oppervlak onder de curve gelijk is aan 1. De functie $f_c(\hat{c}; \xi)$ wordt de *likelihood*-functie van de waarnemingen genoemd bij een willekeurig aangenomen waarde $\rho = \xi$. Het is de waarde van de samengestelde normaal verdeelde kansdichtheidsfunctie van de waarnemingen, waarbij de onderlinge correlaties volgen uit het ruimtelijk statistische model, waarbij $\rho = (1 - \alpha) = \xi$.

We kiezen we nu voor de a priori verdeling $f'_\rho(\xi)$ een zogenaamde 'vage' à priori verdeling, namelijk $f'_\rho(\xi) = 1 (0 \leq \xi < 1)$. We veronderstellen immers dat we geen andere voorinformatie hebben dan dat ρ tussen 1 en nul moet liggen. Dan is de à posteriori schatter $f''_\rho(\xi | \hat{c})$ van de

kansdichtheidsfunctie van ρ gelijk aan de genormaliseerde *likelihood*-functie.

Omdat er niet één unieke waarde van ρ is die er echt uitspringt (zie figuren 2...5 in de hoofdttekst) wordt bij het bepalen van de karakteristieke schatting van de lokale laaggemiddelden de gehele à posteriori verdeling van ρ meegenomen:

$$C_{loc\ av, kar} = \int_0^1 C_{loc\ av, kar | \rho = \xi} f''_\rho(\xi | \hat{c}) d\xi \quad (1.2)$$

De karakteristieke waarde van lokale gemiddelden kan dus worden berekend door sommatie van de producten van de karakteristieke waarden, bij alle aangenomen waarden van ρ en de bijbehorende likelihood-scores.

eigenschappen die per definitie een positieve waarde hebben, is de aanname van een log-normale verdeling veruit te prefereren boven de doorgaans gebruikelijke aanname van een normale kansverdeling. Uit het rekenvoorbeeld blijkt ook dat de lognormale kansverdeling, beter dan een normale kansverdeling, in staat is om hoge uitbijters in een aan de onderkant fysisch begrensde populatie te 'accommoderen'. Karakteristieke ondergrensschattingen werden in het rekenvoorbeeld nauwelijks beïnvloed door (extreem) hoge uitbijters.

In het ruimtelijke model is de verhouding tussen de lokale variantie en de regionale variantie een uiterst belangrijke parameter. Het grondonderzoek, om een regionale proevenverzameling op te bouwen, moet dan ook zodanig opgezet worden dat voldoende informatie over die verhouding ingewonnen wordt. Concreet betekent dit dat proevenverzamelingen voldoende clusters van lokale waarnemingen moeten bevatten. Als stelregel geldt dat een opzet waarbij vier of vijf monsters per boring worden beproefd, te prefereren is boven een opzet met veel meer boringen, maar waarbij slechts één of twee monsters per boring worden beproefd. Omdat de laatste inherent weinig informatie bevatten over de verhouding van lokale en regionale varianties zal, om toch tot veilige schattingen van karakteristieke lokale laaggemiddelden te komen, een 'veilige' waarde van de variantieverhouding moeten

worden aangenomen, zoals bijvoorbeeld in de Leidraad voor het Ontwerpen van Rivierdijken wordt aangegeven.

Het gebruik van regionale proevenverzamelingen voor schuifsterkteparameters van grondlagen is zo'n twintig jaar geleden geïntroduceerd in de Leidraad voor het Ontwerpen van Rivierdijken. Het onderliggende ruimtelijke stochastische model, met bijbehorende statistische formules, was naar de kennis van toen best practice. Dit model is ook wel toegepast bij latere infrastructuurprojecten. Aanbevolen wordt om bij nieuw in te winnen proevenverzamelingen de in dit artikel voorgestelde statistische berekeningsmethode te gebruiken.

Statistische berekeningen op basis van het hier voorgestelde ruimtelijke stochastische model zijn rekentechnisch aanzienlijk complexer dan berekeningen volgens de klassieke statistische steekproeftheorie. Om de berekeningsmethode toegankelijk te maken is een Excel spreadsheet-applicatie ontwikkeld. Dit is op aanvraag (kosteloos) verkrijgbaar bij de auteur. ■

Reacties op dit artikel kunnen tot 1 april 2008 naar de uitgever worden gestuurd.

Referenties

- [1] Calle E.O.F. (2007) *Statistiek bij Proevenverzamelingen*, Geotechniek, juli 2007.
- [2] Vrouwenvelder A. & E. Calle (2003) *Measuring Spatial Correlation of Soil Properties*. Heron, vol. 48 (2003) Nr. 4, pp 297-31.
- [3] TAW (1989) *Leidraad voor het ontwerpen van Rivierdijken* (deel 2). Technische Adviescommissie voor de Waterkering
- [4] CUR (2007) *Van onzekerheid naar veiligheid*. Rapportage CUR commissie C135 (te verschijnen in 2007)
- [5] Ed Calle, Ton Vrouwenvelder, Jaap Lindenberg, Geerhard Hannink en Edward Bruijn (2008) *Representatieve waarden voor grondparameters in de geotechniek* (beoogd te verschijnen in Geotechniek, april 2008)