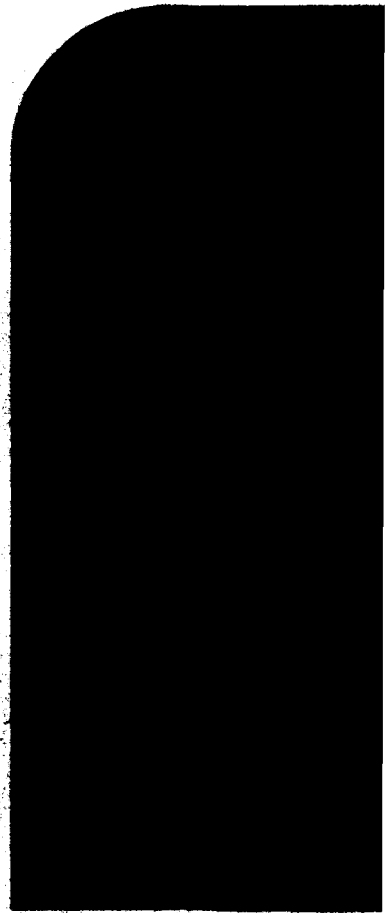
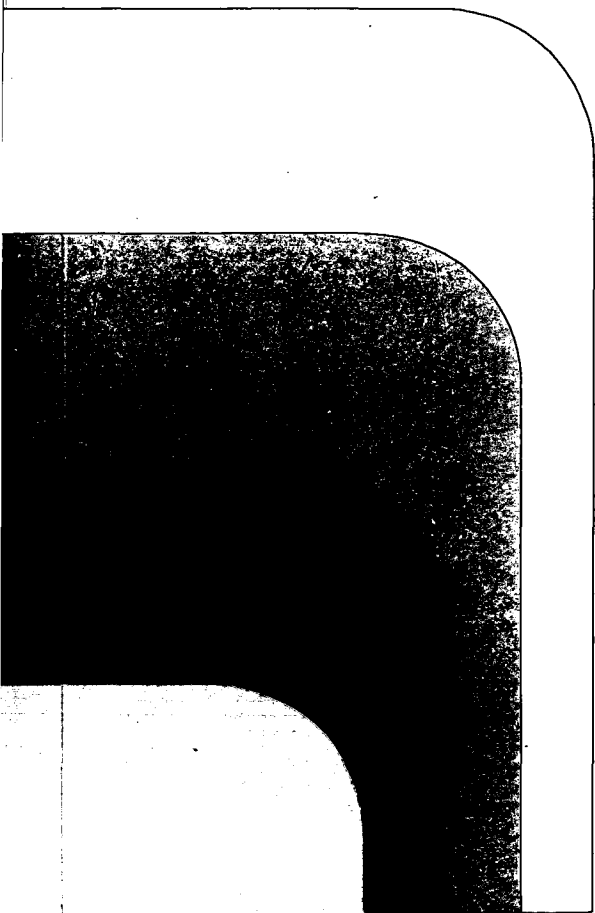




technische adviescommissie voor de waterkeringen



TAW-onderzoeksrapport  
BB87-09

Ruimtelijke autocorrelatie-  
structuur bij grond-  
eigenschappen

E.O.F. Calle  
A.C.W.M. Vrouwenvelder

Rapport C0266483/14  
Grondmechanica Delft  
IBBC-TNO

Sneltjesweg 2  
Postbus 29 2600 AB Delft  
Telefax 015-610621  
Tele 382345011  
Telefoon 015-569223

Postgiro 234342  
Bank Mees en Robbe NV  
Rekening 259235911  
K.v.K. 5145040 Delft



GRONDMECHANICA  
DELFT

RUIMTELIJKE AUTOCORRELATIESTRUCTUUR BIJ GRONDEIGENSCHAPPEN  
(METHODEN VOOR DE ANALYSE VAN VELDMETINGEN)

rapport CO-266483/14

Juli 1987

E.O.F. Calle (Grondmechanica Delft)  
A.C.W.M. Vrouwenvelder (IBBC-TNO)

bladnummer :  
ons kenmerk: C0266483/14  
datum : 1987-07-28



GRONDMECHANICA  
DELFT

INHOUDSOPGAVE

1. INLEIDING	PAG 2
2. DIRECTE PROCEDURES VOOR HET BEREKENEN VAN DE AUTO-CORRELATIEFUNCTIE UIT PUNTWAAARNEMINGEN VAN DE TERREINVARIABLE.	3
3. EEN BAYESIAANSE METHODE VOOR HET SCHATTEN VAN AUTO-CORRELATIEPARAMETERS	16
4. TRENDANALYSE BIJ NORMAAL VERDEELDE STOCHASTISCHE VELDEN	30
5. SAMENVATTING	45
6. REFERENTIES	47

APPENDIX 1: Bewijs dat Tabba & Yong procedure leidt tot zuivere schatters voor de fluctatieschaal.

## 1. INLEIDING EN PROBLEEMSTELLING.

In het kader van de studie 'Probabilistische Stabiliteitsanalyse' zijn methoden onderzocht om de effecten van inhomogeniteit van de grondsterkte eigenschappen in rekening te brengen bij de analyse van stabiliteit van taluds (ref. [1]). Met inhomogeniteit bedoelen we het van plaats tot plaats (binnen een grondlaag) op min of meer grillige wijze (random) variëren van een eigenschap van de grond. Het volledig in kaart brengen van deze variaties zou zeer uitgebreid grondonderzoek vergen, waarbij het variatiepatroon zodanig wordt bemonsterd, dat op een willekeurige plaats tussen de meetpunten de waarde van de eigenschap nauwkeurig kan worden bepaald door interpolatie. In ref. [2],[3] is aangegeven hoe groot de afstanden tussen de meetpunten zouden moeten zijn opdat de interpolatiefout voldoende klein is. Daaruit blijkt dat die afstanden klein tot zeer klein zijn ten opzichte van de in de huidige praktijk van grondonderzoek gehanteerde afstanden. De kosten van grondonderzoek dat zodanig is ingericht dat het variatiepatroon van grondeigenschappen volledig kan worden gereconstrueerd, zullen daarom in het algemeen niet acceptabel worden gevonden. Om deze reden moeten we bij het in rekening brengen van de effecten van ruimtelijke variatie uitgaan van een statistische beschrijving van het variatiepatroon. In de modellen waarin de ruimtelijke variatie van een grondeigenschap wordt meegenomen, zien we dat dit variatiepatroon wordt opgevat als een realisatie van een stochastisch proces (ook wel een stochastisch veld genoemd). Een verdere aanname is dat de kansdichtheidsverdeling van de variabele grootte in elk punt binnen de grondlaag hetzelfde is en dat de onderlinge correlatie tussen de waarden van de variabele in elk tweetal punten binnen de grondlaag uitsluitend een functie van de afstand tussen die punten is of een functie van alleen de afstandscomponenten in de verschillende ruimtelijke dimensies. Zo'n veld wordt een stationair veld genoemd. Deze terminologie komt voort uit de tijdreeksanalyse, waar een stochastisch proces een gebeuren in de tijd is en waar de tijdsonafhankelijkheid van statistische grootheden wordt aangeduid met stationariteit. De processen die we hier beschouwen zijn variaties in een gebied met ruimtelijke dimensies, vandaar de term stochastisch veld. De plaatsonafhankelijkheid van statistische grootheden zullen we in plaats van met stationariteit ook wel met statistische homogeniteit aanduiden. In een variatiepatroon dat continu is zijn punten die heel dicht bij elkaar liggen sterk gecorreleerd. Punten die ver uit elkaar liggen zijn zwak of niet gecorreleerd. De functie die de afname van de correlatie beschrijft als functie van de afstand tussen twee punten heet de autocorrelatiefunctie. De functiewaarde van de autocorrelatiefunctie is gelijk 1.0 als de afstand tussen de punten gelijk nul is (of zeer klein) en zal naar nul gaan als de afstand tussen de punten groot is. Het verloop van de autocorrelatiefunctie is een maat voor de "gemiddelde snelheid" waarmee fluctuaties van de grondeigenschap optreden. Een snel uitdempende autocorrelatiefunctie wijst op snelle fluctuaties, een langzaam uitdempende autocorrelatiefunctie wijst op langzame fluctuaties.

In de praktijk moet de autocorrelatiefunctie worden bepaald uit waarnemingen van de veldgrootte in een (groot) aantal punten. In dit



rapport zullen procedures om de autocorrelatiefunctie te berekenen uit puntwaarnemingen worden beschouwd. Daartoe wordt eerst een opsomming gegeven van de bestaande rekenmethoden, voor zover achterhaalbaar uit de literatuur, en worden deze methoden geanalyseerd. Het doel van die analyse is om vast te stellen met welke nauwkeurigheid de autocorrelatiefunctie volgens de betreffende methode kan worden bepaald in relatie tot het aantal puntwaarnemingen. Omdat tijdens het onderzoek bleek dat de bestaande methoden nogal onbetrouwbare schattingen van de autocorrelatiefunctie opleveren als het aantal puntwaarnemingen niet extreem groot is, is een zogenaamde 'Bayesian inference' procedure ontwikkeld, die, naar het zich laat aanzien, redelijk betrouwbare schattingen oplevert bij een niet al te groot aantal puntwaarnemingen.

Bij het bepalen van de autocorrelatiefunctie uit puntwaarnemingen kan een storend effect een belangrijke rol spelen. Het is namelijk mogelijk dat de waarnemingsprocedure toevallige of systematische fouten bevat. Indien de waarnemingsprocedure toevallige fouten bevat, dat wil zeggen dat bij het waarnemen een fout wordt gemaakt die overigens bij elke waarneming verschillend en onafhankelijk van de fout bij de vorige waarnemingen is, dan zal in een continu variatiepatroon een autocorrelatiefunctie worden gevonden die in het nulpunt een discontinuïteit vertoont. De autocorrelatiefunctie is dan gelijk 1.0 als de afstand tussen de punten gelijk nul is, immers een waarneming is altijd volledig met zichzelf gecorreleerd. Voor elke willekeurig kleine afstand groter dan nul (andere waarneming) is de autocorrelatie een eindige waarde kleiner dan 1.0, terwijl voor grotere afstanden de autocorrelatie monotoon daalt tot nul. In de geostatistische literatuur wordt dit effect het 'nugget effect' genoemd. Indien de waarnemingsprocedure systematische fouten bevat, die voor alle waarnemingen dezelfde is, dan zal dit niet de statistische grootheden van het variatiepatroon beïnvloeden. De schattingen van de gemiddelde waarde en eventueel de gemiddelde trend worden hierdoor wel beïnvloed.

## 2. DIRECTE PROCEDURES VOOR HET BEREKENEN VAN DE AUTOCORRELATIEFUNCTIE UIT PUNTWAAARNEMINGEN VAN DE TERREINVARIABLE.

In dit hoofdstuk zullen we achtereenvolgens een aantal bekende methoden beschrijven, waarmee de autocorrelatiefunctie van een statistisch homogeen variatiepatroon kan worden geschat, op basis van puntwaarnemingen. Voor zover mogelijk zal worden aangegeven wat de betrouwbaarheid van de schattingsmethode is.

### 2.1 Directe berekening bij equidistante puntmetingen op een lijn.

Bij deze methode gaan we uit van de meting van een terreinvareabele in een aantal punten op een rechte lijn. De afstand tussen de opvolgende punten op de lijn is steeds constant, stel  $\Delta x$ . Het variatiepatroon van de veldvariabele op de lijn wordt opgevat als een realisatie van een eendimensionaal stochastisch proces  $w(x)$ . Dit proces wordt gedacht normaal verdeeld te zijn met verwachtingswaarde  $\mu_w = 0$  en

standaardafwijking  $\sigma_\omega = 1$ . Elk stationair normaal verdeeld proces kan zodanig worden getransformeerd dat hieraan voldaan wordt. De autocorrelatiefunctie van het proces geven we aan met  $\rho_\omega(\tau)$ . Deze functie wordt onbekend verondersteld. De meting geven we aan met  $\{\omega_i = \omega(i\Delta x)\}$  ( $i=1 \dots n$ ). We beschouwen de volgende statistische bewerking van de meting:

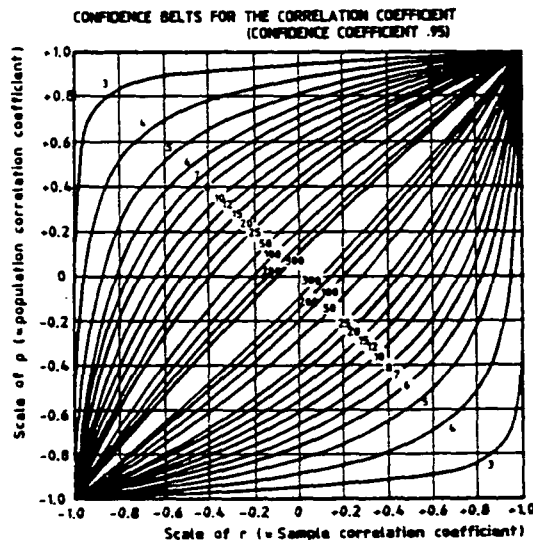
$$r_k = \frac{1}{n-k} \sum_{i=1}^{n-k} \omega_i \omega_{i+k} \quad (k=1, 2, \dots) \quad (2.1)$$

We kunnen de verwachtingswaarde van  $r_k$  bepalen:

$$E[r_k] = \frac{1}{n-k} \sum_{i=1}^{n-k} E[\omega_i \omega_{i+k}] = \frac{\sigma_\omega^2}{n-k} \sum_{i=1}^{n-k} \rho_\omega(k\Delta x) = \rho_\omega(k\Delta x) \quad (2.2)$$

Met  $E[\ ]$  wordt de verwachting aangeduid. De statistische grootte  $r_k$  is dus een zuivere schatter voor de autocorrelatie  $\rho_\omega(k\Delta x)$ . Voeren we de procedure uit voor  $k=1, 2, \dots$ , dan kunnen we een representatie van de autocorrelatiefunctie tekenen in de vorm van een puntendiagram (een zogenaamd correlogram).

In de grafiek in figuur 1 wordt het 95 % betrouwbaarheidsgebied voor de werkelijke waarde van de autocorrelatie  $\rho_\omega(k\Delta x)$  weergegeven als functie van de geschatte waarde  $r_k$  en het aantal producten van waarnemingen waarover gesommeerd wordt ( $n-k$ ). Deze grafiek is overgenomen uit ref. [4].



Figuur 1. Betrouwbaarheid van schattingen van correlatie (overgenomen uit Dixon, ref. [4]).

bladnummer : - 5 -  
 ons kenmerk: C0266483/14  
 datum : 1987-07-28

Zouden we bijvoorbeeld vinden  $r_k=0.80$ , dan geldt dat de werkelijke waarde  $\rho_w(k\Delta x)$  met kans 0.95 ligt tussen de grenzen:

0.75 en 0.83 als  $n-k=400$   
 0.71 en 0.86 als  $n-k=100$   
 0.60 en 0.90 als  $n-k=25$   
 0.30 en 0.95 als  $n-k=10$

We zien dat de kwaliteit van de schatting snel slechter wordt als we niet over grote aantallen puntmetingen beschikken. Is de waarde van de schatter klein, bijvoorbeeld  $r_k=0$ , dan is het 0.95 betrouwbaarheidsgebied nog groter, namelijk:

-0.10 tot 0.10 als  $n-k=400$   
 -0.20 tot 0.20 als  $n-k=100$   
 -0.40 tot 0.40 als  $n-k=25$   
 -0.60 tot 0.60 als  $n-k=10$

Deze resultaten geven aan dat het bepalen van de autocorrelatiefunctie langs deze weg op basis van een beperkt aantal punten een moeizame zaak is.

## 2.2 Uitbreiding van de directe methode voor regelmatig of onregelmatig verspreid liggende meetpunten in een twee-dimensionaal veld.

Beschouw het twee-dimensionale veld van meetpunten in figuur 2. De meetpunten liggen op een regelmatig grid. De gridafstand geven we aan met  $a$ . Het is niet noodzakelijk dat in elk gridpunt een meting gegeven is. We kunnen de methode van de vorige paragraaf toepassen in verschillende richtingen, bijvoorbeeld N-Z (noord-zuid), O-W, NO-ZW of NW-ZO. In N-Z en O-W richtingen vinden we schattingen voor de autocorrelatie op afstanden  $a$ ,  $2a$  enz. en in de NO-ZW en NW-ZO richtingen vinden we schattingen voor de autocorrelatie op afstanden  $a/\sqrt{2}$ ,  $2a/\sqrt{2}$  enz. De aantallen waarnemingen voor de schattingen in de verschillende richtingen zijn:

aantal/ richting:	afstand					
	$a$	$a/\sqrt{2}$	$2a$	$2a/\sqrt{2}$	$3a$	$3a/\sqrt{2}$
N-Z	26		19		12	
O-W	25		21		16	
NO-ZW		21		14		8
NW-ZO		21		15		7





Is er significant verschil in correlatieafval in de verschillende richtingen, dan moeten we concluderen dat de autocorrelatiestructuur anisotroop is. Indien er geen aanwijzingen bestaan om anisotropie te veronderstellen dan kunnen we de aantallen waarnemingsparen met dezelfde ondelinge afstand samenvoegen. We krijgen dan de volgende totale aantallen van waarnemingsparen voor de schatting van de autocorrelatiefunctie:

afstand:	a	a/2	2a	2a/2	3a	3a/2
aantal:	51	42	40	39	28	15

Deze methodiek wordt gepropageerd in de mijngeostatistiek (zie o.a. Journell & Huijbregts ref. [9], Davies ref. [10]). Voor wat betreft de betrouwbaarheid van de schattingen van de autocorrelatie geldt hetzelfde als aangegeven is in de vorige paragraaf.

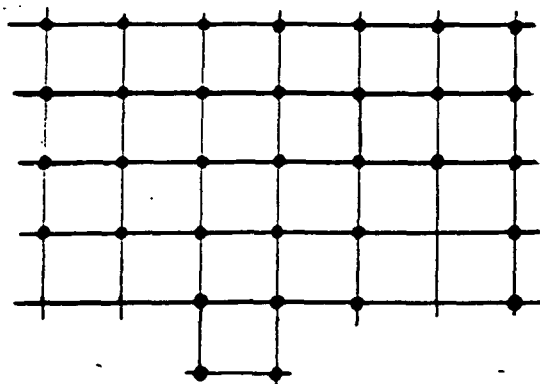
Een verdere modificatie van de methode is nodig als de meetpunten niet op een regelmatig grid liggen (zie o.a. Tang ref. [12]). In dat geval moeten we klassen van afstanden en eventueel klassen van richtingen definiëren. Bijvoorbeeld voor de schatting van  $\rho(ia)$  in N-Z richting worden alle waarnemingsparen gebruikt waarvan de onderlinge afstand tussen  $(i-1/2)a$  en  $(i+1/2)a$  ligt en waarvan de richting van de verbindingslijn tussen de locaties van de waarnemingen die een paar vormen tussen NNO-ZZW en NNW-ZZO ligt. Hebben we te maken met een isotrope correlatiestructuur, dan vervalt natuurlijk het laatste te maken onderscheid.

Opmerking:

In de literatuur, vooral die met betrekking tot de mijngeostatistiek, worden de beschouwingen toegespitst op het zogenaamde semi-variogram, in plaats van op de autocorrelatiefunctie. De semi-variogramwaarde is de halve variantie van het verschil tussen de waarnemingen in twee punten met een bepaalde onderlinge afstand. Deze waarde is, evenals de autocorrelatiefunctie afhankelijk van de afstand. De relatie tussen het semi-variogram en de autocorrelatiefunctie is:

$$\frac{1}{2} \gamma(a) = \sigma^2(\omega) \{1 - \rho_{\omega}(a)\} \quad (2.3)$$

waarin  $\gamma(a)$  de variogramwaarde ( $\frac{1}{2} \gamma(a)$  de semi-variogramwaarde) is. De beschouwingen in deze notitie zijn steeds opgehangen aan de autocorrelatiefunctie, maar kunnen direct vertaald worden in termen van het semi-variogram.



Figuur 2. Waarnemingspunten in een tweedimensionaal veld.

### 2.3 Bepaling van de autocorrelatiefunctie bij continue registraties van een stochastisch veld.

De methoden van de vorige paragrafen waren erop gericht de waarde van de autocorrelatiefunctie te bepalen voor verschillende afstanden tussen de waarnemingspunten. Wanneer we daarbij voor voldoende verschillende afstanden schattingen van de autocorrelatie kunnen bepalen hebben we een indruk hoe de autocorrelatiefunctie van het stochastische veld eruit ziet. In de schattingsprocedures speelde het gegeven dat de autocorrelatiefunctie in het algemeen een uitdempende functie is geen enkele rol. Bij de nu volgende methoden zullen we uitgaan van een functioneel verband tussen autocorrelatie en afstand, waarbij een (of meer) parameter(s) in die expressie moeten worden bepaald.

We gaan uit van een stochastisch veld zoals aangegeven is in paragraaf 2.1 We veronderstellen een expressie voor de autocorrelatiefunctie van de vorm:

$$\rho_{\omega}(\tau) = \exp(-\tau^2/D^2) \quad (2.4)$$

waarin  $D$  een autocorrelatieparameter is. In overeenstemming met Vanmarcke (ref. [5]) kunnen we de zogenaamde fluctuatieschaal ("scale of fluctuation") definiëren als:

$$a = \frac{1}{\rho_{\omega}(0^+)} \int_0^{\infty} \rho_{\omega}(\tau) d\tau \quad (2.5)$$

als deze integraal bestaat, waaruit volgt:  $a = 1/2 D \sqrt{\pi} \approx 0.89 D$  en met behulp hiervan  $\rho_{\omega}(a) \approx 0.54$ . Met de fluctuatieschaal wordt dus die afstand aangegeven waarbinnen een redelijke mate van autocorrelatie en dus een relatief kleine spreiding aanwezig is. Berekend kan worden dat bij een eendimensionaal veld de variantie in een deelgebied met afmeting gelijk aan de fluctuatieschaal de variantie gelijk is aan circa 20 procent van de variantie die in het gehele veld voorkomt,

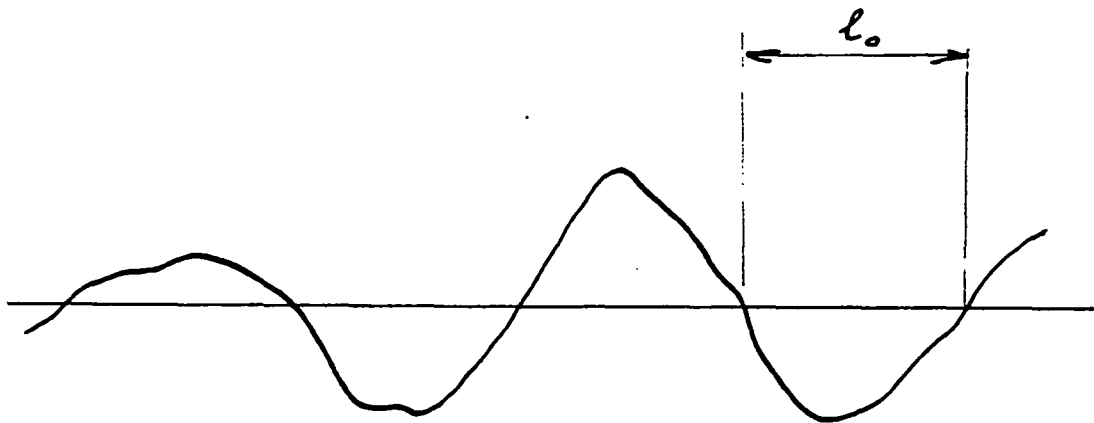
althans bij een autocorrelatiefunctie van de vorm volgens vergelijking (2.4).

We veronderstellen nu dat we over een continue of bijna continue registratie in een raai van het te onderzoeken stochastische veld beschikken (zie figuur 3). De dichtheid van de nuldoorgangen (van positief naar negatief), dat is de verwachtingswaarde van het aantal neerwaartse nuldoorgangen per lengte eenheid, kunnen we voor de hier beschouwde stochastische velden theoretisch berekenen (zie o.a. Papoulis ref. [6] of Klaver ref. [7]):

$$n_0 = \frac{1}{2\pi} \sqrt{-\rho''(0)} = 0.225 D \quad (2.6)$$

waarin  $\rho''(0)$  de tweede afgeleide van de autocorrelatiefunctie is voor  $\tau=0$  en  $n_0$  de dichtheid van de nuldoorgangen is. In formule (2.6) is gerekend met een autocorrelatiefunctie volgen het recept (2.4), dus  $\rho''(0)=-2/D$ . De verwachtingswaarde van de afstand tussen een nuldoorgang van boven naar beneden en de volgende nuldoorgang van beneden naar boven is:

$$E[l_0] = \frac{1}{2n_0} = \frac{\pi D}{\sqrt{2}} = 2.22 D \quad (2.7)$$



Figuur 3. Continue registratie in een raai van een stochastisch veld

We zouden deze lengte kunnen opvatten als de "halve gemiddelde golflengte in het variatiepatroon". In de registratie van het stochastische veld kunnen we de afstanden tussen de nuldoorgangen bepalen en aan de hand hiervan de gemiddelde waarde bepalen. Een schatting van de autocorrelatieparameter  $D$  vinden we door de gevonden

gemiddelde waarde van de afstanden tussen de nuldoorgangen (beter gezegd tussen een nuldoorgang van boven naar beneden tot de opvolgende nuldoorgang van beneden naar boven) in formule (2.7) gelijk te stellen aan  $E[l_0]$ . Om aan te geven wat de nauwkeurigheid van deze methode is zullen we de statistische kenmerken van de afstand tussen twee nuldoorgangen moeten weten. Volgens Vanmarcke (ref. [5]) is deze afstand bij benadering Poisson verdeeld. De kansverdelingsfunctie is:

$$P[l_0 < \xi] = 1 - \exp(-2 n_0 \xi) \quad \text{als } \xi > 0$$

$$= 0 \quad \text{elders} \quad (2.8)$$

Strikt genomen is deze verdeling alleen juist voor doorgangen door relatief hoge niveau's, dus bijvoorbeeld door het  $2\sigma_w$  niveau, maar omdat we hier slechts geïnteresseerd zijn in een globale indicatie van de spreiding van de afstanden tussen nuldoorgangen nemen we de hiermee gemoeide onnauwkeurigheid voor lief. Uit (2.8) vinden we voor de verwachtingswaarde hetzelfde resultaat als (2.6), hetgeen verwacht mocht worden, voor de variantie van  $l_0$  vinden we:

$$\sigma^2(l_0) = E[l_0^2] - E[l_0]^2 = \frac{(\pi D)^2}{2} \quad (2.9)$$

zodat de variatiecoëfficiënt van  $l_0$  is:

$$V_{l_0} = \sigma_{l_0} / E[l_0] = 1 \quad (2.10)$$

In de continue registratie van het stochastische veld kunnen we de afstanden tussen opeenvolgende nuldoorgangen opmeten. Deze afstanden geven we aan met  $\{l_0^i\}$ ,  $i=1,2,\dots,n$ . Het rekenkundig gemiddelde van deze waarden is:

$$\bar{l}_0 = \frac{1}{n} \sum_{i=1}^n l_0^i \quad (2.11)$$

en is een zuivere schatter voor  $E[l_0]$ . We kunnen dus de autocorrelatieparameter  $D$  zuiver schatten met:

$$\tilde{D} = \frac{\bar{l}_0 \sqrt{2}}{\pi} \quad (2.12)$$

waarbij dus:

$$E[\tilde{D}] = D \quad (2.13)$$

De variatiecoëfficiënt van  $\tilde{D}$  is  $1/\sqrt{n}$ . Dit betekent dat we om de waarde van  $D$  nauwkeurig te kunnen schatten door opmeten en middelen van de

afstanden tussen nuldoorgangen in de registratie, we de beschikking zullen moeten hebben over een registratie met een groot aantal nuldoorgangen. Om bijvoorbeeld de waarde van  $D$  met circa 40 procent nauwkeurigheid te schatten, de variatiecoëfficiënt moet dan 0.2 zijn, moeten er 25 nuldoorgangen (van boven naar beneden) in de registratie aanwezig zijn. De totale lengte van de registratie moet dan, gebruik makend van formule (2.5), minimaal 50 a 60  $D$  zijn.

Voor het toepasbaar zijn van deze methode is het nodig te beschikken over een of meer continue registratie(s) van het veld, of over puntmetingen op een lijn die zo dicht bij elkaar liggen dat ze beschouwd kunnen worden als een continue registratie. Dit beperkt de toepasbaarheid van de methode in de praktijk tot die grond- of grondlaageigenschappen die continu of nagenoeg continu gemeten kunnen worden (geo-elektrisch, seismisch, m.b.v. grondradar of via continue sonderen). Verder dienen de registraties zich uit te strekken over gebieden die groot zijn ten opzichte van de "gemiddelde golflengte" in het variatiepatroon.

#### 2.4 Bepaling van de autocorrelatieparameters uit puntmetingen met behulp van curve-fitting.

De methodiek in paragraaf 2.1 is toepasbaar indien de beschikbare puntmetingen zijn uitgevoerd in punten die op een rechte lijn liggen, of in het twee dimensionale geval op een rechthoekig grid liggen. In deze paragraaf zullen we een methode beschouwen waarbij deze beperking kan vervallen. De methode is gepubliceerd door Tappa & Yong (ref. [8]).

We veronderstellen dat een verzameling puntmetingen in een tweedimensionaal veld gegeven is. Die verzameling duiden we aan met  $\{\omega_i = \omega(x_i, y_i)\}$  ( $i=1 \dots n$ ). De methode kan op voor de hand liggende wijze worden uitgebreid naar driedimensionale velden. Het gemeten variatiepatroon wordt weer opgevat als een realisatie van een stationair stochastisch veld. De veldvariabele  $\omega$  wordt weer gedacht op standaardvorm te zijn gebracht, d.w.z. dat de verwachtingswaarde van  $\omega$  is gelijk nul en de standaardafwijking is gelijk 1.0. De autocorrelatie tussen twee realisaties  $\omega_i$  en  $\omega_j$  duiden we aan met  $\rho_{ij}$ , en deze is weer uitsluitend een functie van de afstand tussen de punten  $(x_i, y_i)$  en  $(x_j, y_j)$ , deze afstand noteren we kortweg met  $\tau_{ij}$ . We veronderstellen dat de structuur van de autocorrelatiefunctie bekend is en dat hierin slechts één autocorrelatieparameter voorkomt die moet worden bepaald uit de meting. Die structuur noteren we als:

$$\rho_{ij} = r(D, \tau_{ij}) \quad (2.14)$$

waarin  $D$  de te zoeken autocorrelatieparameter is. De structuur kan bijvoorbeeld zijn:

$$\rho_{ij} = \exp(-\tau_{ij}^2 / D^2) \quad (2.15)$$

of:

$$\rho_{ij} = 1 - \tau_{ij}/D \quad (\tau_{ij} < D) \quad (2.16)$$

Beschouw nu de uitdrukking voor een gekozen  $i$ :

$$R_i(\bar{D}) = \sum_{\substack{j=1 \\ j \neq i}}^n (\omega_i - r(\bar{D}, \tau_{ij}) \omega_j)^2 \quad (2.17)$$

waarin  $\bar{D}$  een schatter voor  $D$  voorstelt.  $R_i(\bar{D})$  willen we minimaliseren onder variatie van  $\bar{D}$ . Een noodzakelijke conditie voor een minimum is dat:

$$\frac{\partial R_i}{\partial \bar{D}} = -2 \sum_{\substack{j=1 \\ j \neq i}}^n (\omega_i - r(\bar{D}, \tau_{ij}) \omega_j) \omega_j \frac{\partial r}{\partial \bar{D}} = 0 \quad (2.18)$$

Als we veronderstellen dat deze vergelijking tot een eenduidig bepaalde waarde van  $\bar{D}$  leidt, dan is de voorwaarde (2.18) voldoende voor het bestaan van een eenduidig minimum. In Appendix 1 wordt aangetoond dat in dat geval de gevonden waarde van  $\bar{D}$  in verwachting gelijk moet zijn aan  $D$ , met andere woorden dat de vergelijking (2.18) een zuivere schatter van  $D$  oplevert.

Tabba & Yong adviseren om voor  $i$  in de vergelijkingen (2.17) - (2.18) die waarneming te kiezen waarbij  $\omega$  in absolute waarde het kleinst is.

Een modificatie van de procedure bestaat daaruit dat bij de residubepaling (2.17) tevens gesommeerd wordt over alle  $i$ , dus:

$$R(\bar{D}) = \sum_{i=1}^n \sum_{j=1}^n (\omega_i - r(\bar{D}, \tau_{ij}) \omega_j)^2 \quad (2.19)$$

De in appendix 1 voor de originele procedure van Tabba & Yong gegeven beschouwing geldt m.m. ook voor de gemodificeerde procedure.

Evenals bij de methoden die in de voorafgaande paragrafen zijn behandeld, bestaat ook hier de mogelijkheid om eventuele anisotropie in de autocorrelatiestructuur te definiëren. Het is dan nodig om in de expressie voor de autocorrelatiefunctie verschillende parameters te introduceren, die de afname van autocorrelatie in de verschillende richtingen in het stochastische veld karakteriseren, bijvoorbeeld  $D_x$ ,  $D_y$  en eventueel  $D_z$ .

Het aangeven van de standaardafwijking van een schatting  $D$  met behulp van deze methode is een ingewikkelde zaak. Toch willen onderzoeken hoe betrouwbaar zo'n schatting is. Bij gebrek aan beter moet dit langs

experimentele weg. Daartoe is een computerprogramma ontwikkeld, waarmee:

- een realisatie van een stochastisch veld kan worden gegenereerd (een zogenaamde pseudo-realisatie), waarbij een autocorrelatiefunctie kan worden opgelegd,
- met behulp van een op bovengenoemde theorie gebaseerde rekenprocedure een schatting van de autocorrelatieparameter (in een eenparametermodel voor de autocorrelatiefunctie) kan worden bepaald door bewerking op het gegenereerde veld,
- deze bewerkingen uit te voeren voor een groot aantal generaties van het veld en gemiddelde en standaardafwijking van de schattingen te bepalen.

In tabel I zijn de uitkomsten gegeven van bewerkingen met behulp van de originele en gemodificeerde Tabba & Yong procedure op pseudo realisaties van eendimensionale velden. Als autocorrelatiefunctie van de gegenereerde velden is het lineaire model (2.16) gekozen. In de tabel zijn de gevonden gemiddelde waarden ( $A_{gem}$ ) en de standaardafwijkingen ( $S_A$ ) van  $A (=1/D^2)$  bij het aangegeven aantal pseudo-realisaties genoemd. De standaardafwijking is een maat voor de betrouwbaarheid van de schatting van  $A$  door bewerking van een pseudo realisatie. Met behulp van de standaardafwijking kunnen we tevens bepalen of de schattingen behept zijn met een significante onzuiverheid. De standaardafwijking van de gemiddelde waarde van  $A$  is namelijk gelijk aan:

$$S_{A_{gem}} = \frac{1}{\sqrt{m}} S_A \quad (2.26)$$

waarin  $m$  het aantal pseudo realisaties is waarover  $A_{gem}$  is berekend. Als  $A$  een zuivere schatter is, dan moet met kans 0.95 gelden:

$$A_{gem} - 2 S_{A_{gem}} \leq A \leq A_{gem} + 2 S_{A_{gem}} \quad (2.27)$$

Hierin is  $A = 1/D^2$ . Wordt hieraan niet voldaan, dan moeten we concluderen dat de schatting van  $A$  onzuiver is.

Tabel I: Numeriek onderzoek van originele en gemodificeerde Tabba & Yong methode, toegepast op een eendimensionaal veld met lineair autocorrelatiemodel.

Gegenereerde puntmetingen op een lijn:  $\{\omega(i\Delta x)\}$  ( $i=1\dots n$ )  
 Karakteristieken van het veld:  $E[\omega]=0$ ,  $\sigma_\omega=1$  en  $\rho_\omega(\tau)=1-A|\tau|$

$n$ <sup>1)</sup>	$\Delta x$ <sup>2)</sup>	Proc. <sup>3)</sup>	$m$ <sup>4)</sup>	A	$A_{gem}$	$S_A$
10	1.0	1	1000	0.05	0.069	0.130
		1	1000	0.05	0.075	0.132
		2	1000	0.05	0.066	0.040
		1	1000	0.10	0.116	0.131
		1	1000	0.10	0.148	0.097
		2	1000	0.10	0.129	0.084
20	0.5	1	100	0.10	0.260	0.200
		2	100	0.10	0.188	0.184
	1.0	1	1000	0.05	0.074	0.133
		1	1000	0.05	0.057	0.061
50	0.5	2	1000	0.05	0.068	0.041
		1	100	0.10	0.139	0.093
		2	100	0.10	0.124	0.084
	0.1	1	100	0.10	0.270	0.216
		2	100	0.10	0.199	0.172
		2	100	0.10	0.206	0.161
0.2	1	100	0.10	0.152	0.106	
	2	100	0.10	0.124	0.084	

- 1) aantal gegenereerde punten in een pseudo realisatie  
 2) onderlinge afstand tussen de punten  
 3) rekenprocedure: 1 = originele, 2 = gemodificeerde Tabba & Yong  
 4) aantal gegenereerde pseudo realisaties

Deze berekeningsresultaten zijn niet erg bemoedigend. In eerste instantie valt op dat de standaardafwijkingen van de schattingen vaak net zo groot of nog groter zijn dan de schattingen zelf. Daarnaast valt te constateren dat de schattingen allesbehalve zuiver zijn. Dit laatste sterker naarmate de correlatieafval in het gebied van de pseudo realisatie betrekkelijk gering is.

In tabel II zijn de resultaten gegeven van berekeningen op gegenereerde eendimensionale velden met een exponentieel afnemende autocorrelatie.



Tabel II. Numeriek onderzoek van originele en gemodificeerde Tabba & Yong procedure op velden met exponentieel afnemende autocorrelatiefunctie.

---

Gegenereerde puntmetingen op een lijn:  $\{\omega(i\Delta x)\}$  ( $i=1\dots n$ )  
 Karakteristieken van het veld:  $E[\omega]=0$ ,  $\sigma_\omega=1$  en  $\rho_\omega(\tau)=\exp(-A\tau^2)$

---

$n$ <sup>1)</sup>	$\Delta x$ <sup>2)</sup>	Proc. <sup>3)</sup>	$m$ <sup>4)</sup>	A	$A_{\text{gem}}$	$S_A$
11	0.2	1	73	1.0	36.3	32.9
		2	98	1.0	1.84	2.07
11	0.5	1	44	1.0	9.84	4.66
		2	98	1.0	1.27	1.08
11	1.0	1	40	1.0	2.67	1.07
		2	89	1.0	0.90	0.67
21	0.2	1	31	1.0	31.6	28.8
		2	99	1.0	1.81	1.78
	0.5	1	21	1.0	7.40	3.43
		2	57	1.0	1.58	1.83
	1.0	1	25	1.0	2.52	0.85
		2	30	1.0	1.25	1.41

---

<sup>1)</sup> aantal punten in het veld

<sup>2)</sup> onderlinge afstand tussen de punten

<sup>3)</sup> aantal simulaties; getracht werd 100 simulaties te realiseren. Het genoemde aantal betreft het aantal simulaties waarbij de iteratieve procedure convergeerde binnen 20 stappen.

<sup>4)</sup> bij de generatie opgelegde waarde van de autocorrelatieparameter A (N.b. A komt overeen met  $1/D^2$  in vergelijking (2.4))

---

Ook bij de resultaten in tabel II hetzelfde beeld als bij de resultaten van tabel I. De uitkomsten van de originele Tabba & Yong procedure zijn soms volstrekt irrelevant, maar ook bij de iets beter gelukte berekeningen zijn ze niet bevredigend.

In tabel III zijn de resultaten opgenomen van berekeningen op pseudo realisaties van twee dimensionale velden. Hierbij is alleen met de gemodificeerde Tabba & Yong procedure gewerkt.

Tabel III. Numeriek onderzoek van de gemodificeerde Tabba & Yong procedure op tweedimensionale velden.

Gegenereerd veld:  $\{\omega(i\Delta x, j\Delta y)\}$  ( $i=1\dots n, j=1\dots n$ )

Karakteristieken van het veld:  $\mu_\omega=0, \sigma_\omega=1$

Autocorrelatiefunctie:  $\rho_\omega(\tau) = \exp(-A\tau^2)$  met  $A = 1.0$

geval	$n \times n$	$\Delta x, \Delta y$	m	$A_{gem}$	$S_A$
1	5x5	1.0	20	1.12	0.60
			50	1.09	0.66
2	7x7	1.0	20	1.03	0.36
			50	0.99	0.42
3	5x5	0.5	20	1.48	1.04
			50	1.44	1.22
4	7x7	0.5	50	1.17	0.62
5	5x5	3.0	50	0.56	0.47

Autocorrelatiefunctie:  $\rho_\omega(\tau) = 1 - A|\tau|$ , met  $A = 0.10$

geval	$n \times n$	$\Delta x, \Delta y$	m	$A_{gem}$	$S_A$
6	5x5	1.0	100	0.16	0.11
7	7x7	1.0	100	0.14	0.07
8	9x9	1.0	100	0.12	0.05
9	10x10	1.11	100	0.11	0.03

De berekeningen voor een tweedimensionaal veld komen gunstiger uit dan de berekeningen voor een lijnmeting. De volgende gevallen bij de berekeningen in tabel III kunnen worden onderscheiden.

- a. De pseudo metingen bevatten informatie over zowel het "support" als over de "staart" van de autocorrelatiefunctie. Onder het support wordt verstaan die afstand op de  $\tau$ -as waar  $\rho_\omega(\tau)$  significant groter dan nul is. Onder de staart verstaan we dat deel van de  $\tau$ -as waar de correlatie nagenoeg nul is. Dit is het geval in de rekegevallen 1, 2, 4 en 9.

- b. De pseudo metingen bevatten wel informatie over een gedeelte van het support, maar niet over de staart (gevallen 3 en 6).
- c. De pseudo metingen bevatten wel informatie over de staart, maar niet over het support (geval 5).

Bij de onder a. genoemde gevallen wordt de naar verhouding beste schatting van de autocorrelatieparameter verkregen. De betrouwbaarheid neemt kennelijk toe naarmate de meting meer informatie over de staart van de autocorrelatiefunctie bevat (geval 2 versus geval 1 en geval 9 versus de gevallen 8, 7 en 6). Bij de gevallen onder b. genoemd bevat de meting voldoende informatie in het support van de autocorrelatiefunctie, maar het ontbreken van voldoende informatie over de staart zorgt ervoor dat de schatting onbetrouwbaar is. Is er geen informatie over het support maar wel over de staart, dan leidt dit ook tot onbetrouwbare schattingen. Dit laatste is gevoelsmatig verklaarbaar, immers bij zulke metingen kan van het support weinig anders worden gezegd dan dat deze kleiner is dan de kleinste onderlinge afstand tussen de meetpunten. De precieze grootte is echter een volslagen onbekende grootte.

## 2.5 Conclusie

Uit de beschouwingen blijkt in elk geval dat de betrouwbaarheid van de schattingen, gebaseerd op slechts een (pseudo) realisatie van het stochastische veld (dus slechts een meting) te wensen overlaat. Belangrijk voor een redelijk betrouwbare schatting van de autocorrelatie structuur lijkt te zijn dat in elk geval dat er een groot aantal waarnemingen in het te onderzoeken veld aanwezig zijn en dat deze zodanig over het veld verspreid liggen dat zowel het support als de staart van de autocorrelatiefunctie door de meting worden gedekt. Voor de directe methode van par. 2.1 is dit in eerste instantie niet strict nodig als het erom gaat schattingen voor de autocorrelatie voor bepaalde afstanden te bepalen. Wanneer uit deze puntschattingen een autocorrelatiefunctie moet worden bepaald, door fitting van een kromme door de gevonden puntwaarden, dan is informatie over het gehele gebied waar de autocorrelatie volledig uitdempt. Beschikken we over een (bijna) continue registratie dan is de methode van par. 2.2 het meest aangewezen. We kunnen hierbij bijvoorbeeld denken aan het vaststellen van de autocorrelatie-afval-functie in verticale richting van sondeerweerstand.

## 3. EEN BAYESIAANSE METHODE VOOR HET SCHATTEN VAN AUTOCORRELATIE-PARAMETERS.

### 3.1 Beschrijving van de methode.

Om autocorrelatieparameters uit een relatief klein aantal waarnemingen te bepalen is een Bayesiaanse schattingsprocedure ontwikkeld. In grote lijnen komt de werking van de procedure op het volgende neer. Veronderstellen we dat de gemeten waarden worden voortgebracht door meting

aan een normaal verdeeld homogeen stochastisch veld met een bepaalde autocorrelatiestructuur, dan is de verzameling meetwaarden een verzameling van gecorreleerde stochasten. De onderlinge correlatie van deze stochasten is alleen afhankelijk van de afstand of afstandscomponenten tussen de meetpunten. De gezamenlijke kansdichtheidsfunctie van deze stochasten is een zogenaamde multivariate normale verdeling:

$$f(\xi_1, \xi_2, \dots; \Omega) = f(\underline{\xi}; \Omega) = \frac{1}{(2\pi)^{n/2} \sigma(\omega)^n |\Omega|^{1/2}} \exp(-1/2 \underline{\theta}^T \Omega^{-1} \underline{\theta}) \dots (3.1)$$

waarin  $n$  het aantal meetpunten is,  $\sigma(\omega)$  de standaardafwijking van het veld,  $\Omega$  een matrix met correlatiecoëfficiënten tussen de waarnemingen:  $\Omega = [\rho_{ij}]$  ( $n \times n$  matrix),  $|\Omega|$  de determinant van de matrix  $\Omega$  en  $\underline{\theta}$  een vector met gestandariseerde stochasten:

$$\underline{\theta} = \left[ \frac{(\xi_1 - \mu(\omega))}{\sigma(\omega)}, \frac{(\xi_2 - \mu(\omega))}{\sigma(\omega)}, \dots \right]^T$$

De correlatiecoëfficiënten  $\rho_{ij}$  die in de matrix  $\Omega$  staan zijn in feite de onbekende grootheden die we willen bepalen. Evenals in (2.14) veronderstellen we een functioneel verband tussen deze coëfficiënten en de afstand tussen de betreffende meetpunten, dus een formele expressie voor de autocorrelatiefunctie, bijvoorbeeld met als enige parameter de fluctuatieschaal  $D$ :

$$\rho_{ij} = r(D; \tau_{ij})$$

waarin  $\tau_{ij}$  de afstand tussen de meetpunten is.

Uitgangspunt voor de Bayesiaanse procedure om de fluctuatieschaal  $D$  te schatten is het theorema van Bayes:

$$P[A_k | B] = P[B | A_k] P[A_k] / P[B] \quad (3.2)$$

waarin  $B$  en  $A_k$  gebeurtenissen zijn, de rij  $A_k$  ( $k=1 \dots n$ ) een volledig stelsel van disjuncte gebeurtenissen, dat wil zeggen dat de doorsnede van elk tweetal een onmogelijke gebeurtenis is (ze sluiten elkaar twee aan twee uit) en dat de vereniging van alle een zekere gebeurtenis is (gebeurtenis die met kans 1.0 optreedt). Voor het bewijs van deze elementaire stelling wordt verwezen naar de literatuur (o.a. ref. [6]). Voor de schattingsprocedure nemen we aan dat de fluctuatieschaal  $D$  precies een van een discreet aantal mogelijke waarden  $D_k$  kan aannemen, waarbij de verzameling  $\{D_k\}$  een verzameling equidistante punten is. We kunnen, gebruik makend van het theorema van Bayes, de kans dat  $D=D_k$  berekenen, gegeven de waarnemingen  $\underline{\omega} = \underline{\xi}$ . Deze kans kunnen we schrijven als:

$$P[D=D_k | \underline{\omega}=\underline{\xi}] = P[\underline{\omega}=\underline{\xi} | D=D_k] P[D=D_k] / P[\underline{\omega}=\underline{\xi}] \quad (3.3)$$

De kans  $P[\underline{\omega}=\underline{\xi} | D=D_k]$  is evenredig met de waarde van de kansdichtheidsfunctie (3.1) waarin de meetwaarden  $\underline{\xi}$  worden gesubstitueerd en waarin de correlatie matrix  $\Omega$  wordt bepaald op basis van de autocorrelatiefunctie met daarin voor  $D$  de waarde  $D_k$ . Deze waarde van de kansdichtheidsfunctie noemen we de likelihoodscore van de waarnemingen bij de aangenomen correlatieparameter. We vinden dus:

$$P[D=D_k | \underline{\omega}=\underline{\xi}] = C f_{\underline{\omega}}(\underline{\xi}; \Omega(D_k)) P[D=D_k] / P[\underline{\omega}=\underline{\xi}] \quad (3.4)$$

De kans  $P[D=D_k]$  in (3.4) noemen we de a priori kans dat  $D$  gelijk is aan  $D_k$ , dat wil zeggen de kans die we hieraan toekennen voordat de actuele waarnemingen  $\underline{\xi}$  bekend zijn. Als verondersteld wordt dat a priori geen van de  $D_k$  waarden kansrijker is dan de overige, is de kans  $P[D=D_k]$  in (3.4) onafhankelijk van  $k$ . De noemer  $P[\underline{\omega}=\underline{\xi}]$  is eveneens onafhankelijk van  $k$ . Beide kansen in het rechterlid van (3.4) kunnen samen met de factor  $C$  worden gecombineerd tot een nieuwe, van  $k$  onafhankelijke, factor, die zodanig moet worden gekozen dat de som van de kansen  $P[D=D_k | \underline{\omega}=\underline{\xi}]$  voor  $k=1 \dots n$  gelijk is aan 1.0. Daarmee hebben we dan, zij het in gediscretiseerde vorm, de zogenaamde a posteriori kansverdeling voor  $D$  gevonden. Met behulp hiervan kunnen we dan bepalen:

- De  $D$  waarde met de hoogste likelihoodscore (maximum likelihood estimate, genoteerd als  $\hat{D}$ ),
- de verwachtingswaarde van  $D$  (notatie:  $E[D]$ ) en
- de standaardafwijking van  $D$  (notatie:  $\sigma_D$ ).

De procedure is rechttoe rechtaan uit te breiden voor autocorrelatiefuncties met meer dan een parameter. In de volgende paragrafen worden de resultaten die met behulp van deze methode worden gevonden nader geanalyseerd.

### 3.2 Resultaten van analyse van gegenereerde velden.

In eerste instantie is, net zoals in hoofdstuk 2 (par. 2.4), gekeken naar gegenereerde een en tweedimensionale velden. De veldvariabele is normaal verdeeld, met verwachtingswaarde gelijk nul en standaardafwijking gelijk 1.0. De opgelegde autocorrelatiefunctie is in alle gevallen van de in vergelijking (2.15) gegeven exponentieel afnemende vorm.

Tabel IV. Berekeningsresultaten van Bayesiaanse procedure bij pseudo realisaties van een eendimensionaal veld.

Gegenereerde puntmetingen op een lijn:  $\{\omega(i\Delta x)\}$  ( $i=1\dots n$ )  
 Karakteristieken van het veld:  $E[\omega]=0$ ,  $\sigma_\omega=1$  en  $\rho_\omega(\tau)=\exp(-\tau^2/D^2)$

$n$ <sup>1)</sup>	$\Delta x$ <sup>2)</sup>	$D$ <sup>3)</sup>	geval	$\hat{D}$ <sup>4)</sup>	$E[D]$ <sup>5)</sup>	$\sigma_D$ <sup>6)</sup>
11	0.2	1.0	1	1.0	0.992	0.035
			2	1.0	0.999	0.030
			3	0.9	0.991	0.029
			4	1.0	1.036	0.051
			5	1.0	1.019	0.044
			gg	1.1	1.095	0.032
			kg	0.9	0.906	0.030
			gs	1.1	1.062	0.052
			ks	1.0	0.997	0.018
			ga	0.9	0.906	0.030
11	0.5	1.0	1	0.9	0.921	0.081
			2	1.1	1.048	0.083
			3	1.0	0.955	0.078
			4	1.0	1.019	0.085
			5	1.1	1.081	0.082
			gg	1.1	1.176	0.090
			kg	0.9	0.821	0.105
			gs	1.0	0.890	0.138
			ks	1.0	1.020	0.070
			ga	1.1	1.140	0.073
11	1.0	1.0	1	0.9	0.844	0.198
			2	1.3	1.158	0.296
			3	1.0	0.904	0.264
			4	1.1	0.826	0.323
			5	1.3	1.111	0.292
			gg	1.3	1.335	0.183
			kg	0.5	0.587	0.229
			gs	1.2	0.994	0.343
			ks	1.3	1.335	0.183
			ga	0.5	0.587	0.229

1) aantal waarnemingen in pseudo realisatie

2) afstand tussen de meetpunten

3) in simulatie opgelegde autocorrelatieparameter

4) waarde van D met hoogste likelihood score

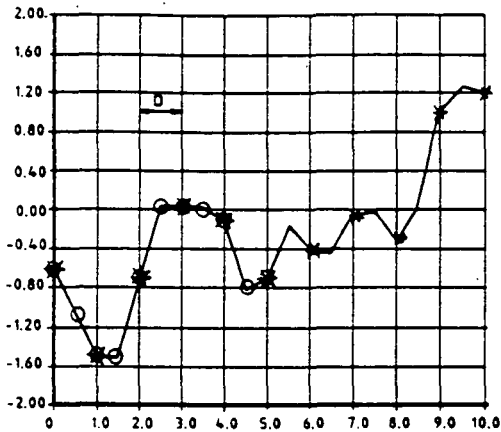
5) verwachtingswaarde van D, berekend uit a posteriori verdeling

6) standaardafwijking van D, berekend uit a posteriori verdeling

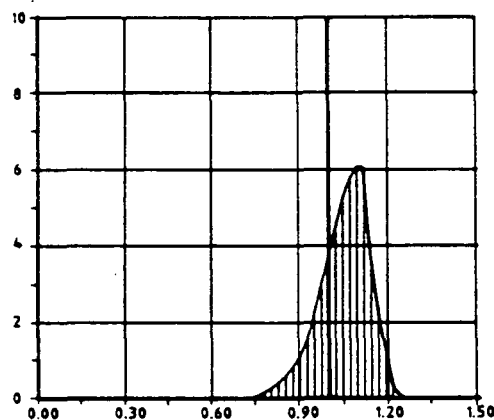


In tabel IV zijn de resultaten gegeven van een aantal berekeningen op gegenereerde eendimensionale velden. Daarbij zijn pseudo realisaties van een elftal punten op een lijn gegenereerd. De onderlinge afstand tussen opvolgende punten is constant. Drie gevallen zijn onderzocht, namelijk de afstanden  $\Delta x=0.2$ ,  $\Delta x=0.5$  en  $\Delta x=1.0$ , waarbij steeds met dezelfde autocorrelatiefunctie (volgens vergelijking 2.15) met fluctuatieschaal  $D=1.0$  is gerekend. Voor elk van die gevallen zijn honderd pseudo realisaties gegenereerd. De in de tabel gegeven uitkomsten betreffen de resultaten van de eerste vijf berekeningen van elke serie (aangeduid met de geval nummers 1 t/m 5) en van elke serie de gevallen waarbij de grootste verwachtingswaarde (gg), kleinste verwachtingswaarde (kg), grootste standaardafwijking (gs), kleinste standaardafwijking (ks) en grootste relatieve afwijking (ga) is gevonden. Onder de grootste relatieve afwijking wordt verstaan de grootste verhouding van het absolute verschil tussen de bij een pseudo realisatie gevonden verwachtingswaarde en 1.0 en de gevonden standaardafwijking.

Bij iedere berekening vinden we een verdeling van de likelihood scores van de verschillende waarden van  $D$ . Deze verdeling noemen we de a posteriori waarschijnlijkheidsverdeling van  $D$ , gegeven de pseudo meting. In figuur 4 zijn de gevonden verdelingen voor een aantal berekeningen weergegeven. Opvallend is dat de spreiding in de verdelingen relatief groot is bij grote afstanden tussen de waarnemingspunten. Bij de afstand  $\Delta x=1$  zien we dat de verdeling van de likelihoodscores naar  $D=0$  toe nauwelijks uitdempt. Dit betekent dat het voor de procedure kennelijk moeilijk wordt de juiste waarde van  $D$  te vinden naarmate de correlatie tussen de dichtst bij elkaar gelegen punten relatief klein is, dat wil zeggen dat de meting feitelijk geen informatie bevat over het autocorrelatie-afval-gedrag in de buurt van  $\tau=0$ . We zien overigens dat de gevonden resultaten aanmerkelijk beter zijn dan die welke gevonden zijn met de directe procedures uit paragraaf 2.4.

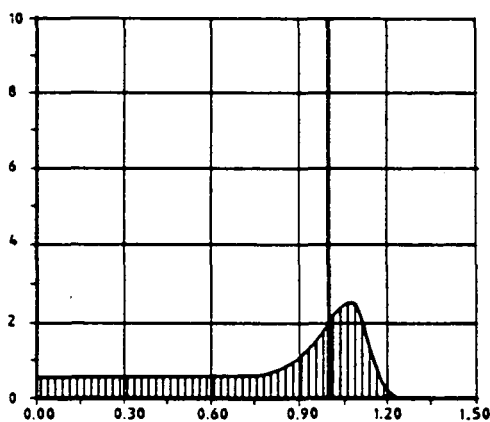


(a): Voorbeeld van een random gegenereerde functie met gemiddelde 0 en standaardafwijking 1 en  $D=1.0$  m. De punten aangegeven met 0 zijn gesampled voor de schatting onder (b) (met  $\Delta x=0.5$  m). De punten aangegeven met \* zijn gesampled voor de schatting onder (c) (met  $\Delta x=1.0$  m).



(b): Kansdichtheidsfunctie van D op basis van 11 waarnemingen met  $\Delta x=0.5$  m.

Verwachting  $E[D]=1.048$  m, standaardafwijking  $\sigma_D=0.083$  m. Goed resultaat (tabel IV,  $\Delta x=0.5$ , geval 2)



(c): Kansdichtheidsfunctie van D op basis van 11 waarnemingen met  $\Delta x=1.0$  m.

Verwachting  $E[D]=0.826$  m, standaardafwijking  $\sigma_D=0.323$  m. Slecht resultaat. Omdat  $\Delta x$  te groot is bevatten de waarnemingen onvoldoende informatie om lage waarden van D uit te sluiten. (tabel IV,  $\Delta x=1.0$ , geval 4)

Figuur 4. Verdeling van likelihoodscores bij de 1-D velden.



In tabel V zijn de resultaten vermeld van soortgelijke exercities met tweedimensionale velden. Hier zien we dezelfde tendenzen. In het eerste geval, met  $\Delta x = \Delta y = 1$  en  $D = 2$  (vergelijk met de gevallen  $\Delta x = 0.5$  en  $D = 1$  in tabel IV), zien we een zeer nauwkeurige schatting van de autocorrelatie parameter. De schattingen zijn nog nauwkeuriger dan de schattingen bij het eendimensionale veld, omdat de pseudo metingen in het tweedimensionale geval veel meer puntwaarnemingen bevatten (36 i.p.v. 11). In het tweede geval, met  $\Delta x = \Delta y = 1$  en  $D = 1$ , zien we overigens hetzelfde effect als bij de eendimensionale berekening ten aanzien van de verdeling van likelihoodscores voor kleine waarden van  $D$ . Ook in het tweedimensionale geval bevat de meting weinig informatie omtrent de autocorrelatie-afval voor kleine waarden van de afstand. Het feit dat in het tweedimensionale geval meer puntwaarnemingen voorhanden zijn doet daar weinig aan af. Kennelijk is het zo dat meer puntwaarnemingen alleen een betere schatting opleveren als in die waarnemingen ook de gezochte informatie is vervat.

Tabel V. Resultaten van Bayesiaanse procedure bij tweedimensionale velden.

---

Gegenereerde puntmetingen op een lijn:  $\{\omega(i\Delta x)\}$  ( $i=1\dots n$ )

Karakteristieken van het veld:  $E[\omega]=0$ ,  $\sigma_\omega=1$  en  $\rho_\omega(\tau)=\exp(-\tau^2/D^2)$

---

$n \times n$ <sup>1)</sup>	$\Delta x = \Delta y$ <sup>2)</sup>	$D$ <sup>3)</sup>	run nr.	$\hat{D}$ <sup>4)</sup>	$E[D]$ <sup>5)</sup>	$\sigma_D$ <sup>6)</sup>
6x6	1.0	2.0	1	2.0	1.999	0.019
			2	2.0	1.995	0.033
6x6	1.0	1.0	1	0.875	0.735	0.228
			2	0.875	0.684	0.218

---

- 1) aantal waarnemingen in pseudo realisatie  
 2) afstand tussen de meetpunten  
 3) in simulatie opgelegde autocorrelatieparameter  
 4) waarde van  $D$  met hoogste likelihood score  
 5) verwachtingswaarde van  $D$ , berekend uit verdeling van likelihood scores  
 6) standaardafwijking van  $D$ , berekend uit verdeling van likelihood scores
-

### 3.3 Resultaten van berekeningen aan veldmetingen.

Ten behoeve van een vergelijkend sondeonderzoek is een meetcampagne uitgevoerd in de jaren 1966 en 1967 op een terrein in Leidschendam, waar later het koopcentrum Leidschenhage is gebouwd. De oorspronkelijke opzet van de meting was om de resultaten van sonderingen met verschillende conusvormen te vergelijken. Het terrein werd gekozen vanwege de verwachte homogene grondslag (diepe zandlaag). De sonderingen zijn verricht op 18 verspreid over het terrein liggende locaties (zie figuur 5). Op elke locatie werden twee sonderingen uitgevoerd met acht verschillende conustypen. Voor ons onderzoek beperken we ons tot de resultaten van twee conustypen, namelijk de zogenaamde mechanische mantelconus en de gladde elektrische conus. Overigens is alleen de laatste heden ten dage in gebruik.

Als veldvariabele wordt voor het correlatieonderzoek de gemiddelde conusweerstand  $\bar{q}_c$  over een diepte van 2.8 m in de draagkrachtige zandlaag (in tabel VI genoteerd in kgf/cm<sup>2</sup>) genomen.

Voor de analyse van de demping van autocorrelatie van deze veldvariabele wordt een gedrag verondersteld van de vorm:

$$\rho(\tau) = a_0 \exp(-\tau^2/D^2)$$

Indien  $a_0$  ongelijk 1.0 blijkt te zijn doen moeten we concluderen dat er een directe afval voor autocorrelatie is voor  $\tau=0^+$ , dat wil zeggen dat metingen, hoe dicht ook bij elkaar onderling zullen verschillen. De variantie van deze verschillen is een maat voor de niet reproduceerbaarheid van de sonderingen. De variantie van de niet reproduceerbaarheid (de reproduceerbaarheidsfout) is gelijk aan  $(1-a_0)$  keer de variantie in de populatie van de meetresultaten. Dit betreft niet systematische meetfouten of variaties op kleine schaal ("small scale fluctuations") of een combinatie van beide. Het overige deel van de variantie, dus  $a_0$  keer de variantie in de populatie van meetresultaten moet worden toegeschreven aan echte terreinvarianties.

De resultaten van de analyse zijn weergegeven in tabel VII.

De verdelingen van likelihoodscores, in dit geval tweedimensionale waarschijnlijkheidsverdelingen, zijn aangegeven in de tabellen VIII en IX.

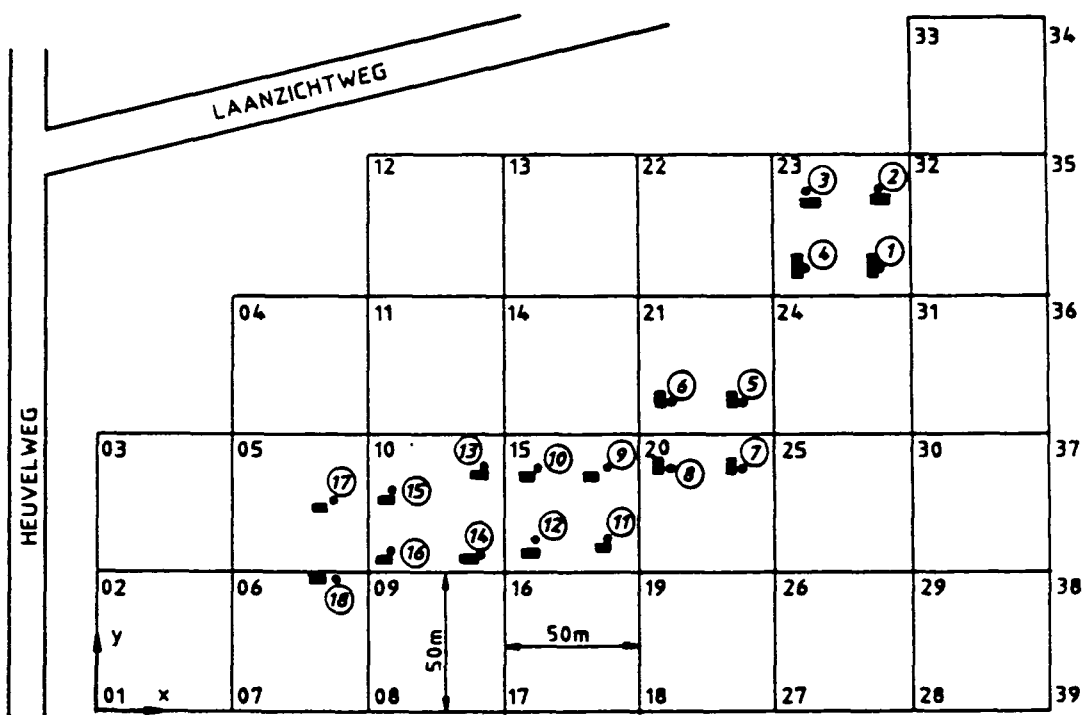
De variatiecoëfficiënt (= standaardafwijking gedeeld door gemiddelde waarde) van de sonderingen bedraagt voor zowel voor de mechanische als voor de elektrische sonderingen ongeveer 0.20 (0.203 voor de mechanische en 0.199 voor de elektrische conus). Rekenend met de gevonden verwachte  $a_0$  waarde van 0.733 voor de mechanische mantelconus vinden we de variatiecoëfficiënt van de echte terreinvarianties:

$$\frac{(a_0 \sigma^2(\bar{q}_c))^{1/2}}{E[\bar{q}_c]} = 0.175$$

en de variatiecoëfficiënt voor meetfout en/of kleine schaal variaties:

$$\frac{((1-a_0) \sigma^2(\bar{q}_c))^{1/2}}{E[q_c]} = 0.105$$

De overeenkomstige waarden voor de elektrische sonderingen zijn 0.185 en 0.075. Het lijkt er dus op dat de sonderingen met de gladde elektrische conus iets beter reproduceerbaar zijn, maar dit is voor wat betreft het in dit rapport aan de orde zijnde problematiek slechts een zijdelingse opmerking. De berekende fluctuatieschalen vertonen elk een aanzienlijke spreiding. Globaal zou men kunnen zeggen dat het invloedsgebied van een sondering ca. 10 a 20 meter is.



Figuur 5. Situatie van proefvakken in het proefterrein.

bladnummer : - 25 -  
 ons kenmerk: C0266483/14  
 datum : 1987-07-28



Tabel VI. Waargenomen gemiddelde conusweerstand in het proefterrein.

MECHANISCHE MANTELCONUS SONDERINGEN:

Sondering no.	x-coord. [m]	y-coord. [m]	$\bar{q}_c$ [kgf/cm <sup>2</sup> ] (=0.1 MPa)	$\frac{\bar{q}_c - E[\bar{q}_c]}{\sigma(\bar{q}_c)}$
1	285.25	158.80	105	1.452
2	286.75	158.80	102	-1.270
3	283.90	188.25	74	-0.430
4	283.90	186.75	60	-1.280
5	258.90	160.75	99	1.088
6	266.10	159.25	85	0.238
7	261.75	191.10	79	-0.127
8	260.25	188.70	68	-0.795
9	238.25	116.10	89	0.481
10	238.25	111.30	82	0.056
11	210.25	116.10	91	0.602
12	214.75	108.90	105	1.452
13	238.25	91.10	86	0.299
14	235.25	83.90	61	-1.220
15	213.25	88.70	98	1.027
16	211.25	83.90	85	0.238
17	191.10	88.25	64	-1.037
18	191.10	85.25	63	-1.098
19	163.70	88.25	91	0.602
20	158.90	86.75	108	1.635
21	191.10	63.25	57	-1.463
22	186.30	60.25	75	-0.369
23	158.90	64.75	76	-0.309
24	161.30	64.75	71	-0.612
25	143.20	88.25	89	0.481
26	143.20	86.75	88	0.420
27	138.40	56.75	81	-0.005
28	140.80	56.75	65	-0.977
29	104.40	92.25	73	-0.491
30	109.20	92.25	82	0.056
31	104.40	57.25	98	1.027
32	106.80	55.75	105	1.452
33	83.90	77.25	60	-1.280
34	91.10	72.75	39	-2.556
35	83.90	50.25	70	-0.673
36	88.70	50.25	95	0.845

Gemiddelde waarde  $E[\bar{q}_c] = 81.08$

Standaarddeviatie  $\sigma(\bar{q}_c) = 16.46 \text{ kgf/cm}^2$

bladnummer : - 26 -  
 ons kenmerk: C0266483/14  
 datum : 1987-07-28



Tabel VI. Vervolg.

GLADDE ELEKTRISCHE CONUS SONDERINGEN:

Sondering no.	x-coord. [m]	y-coord. [m]	$\bar{q}_c$ [kgf/cm <sup>2</sup> ] (=0.1 MPa)	$\frac{\bar{q}_c - E[\bar{q}_c]}{\sigma(\bar{q}_c)}$
1	285.25	163.60	112	0.411
2	289.75	158.80	123	0.943
3	286.30	189.75	109	0.266
4	291.10	188.25	105	0.073
5	263.70	160.75	131	1.331
6	258.90	157.75	125	1.040
7	264.75	191.10	117	0.653
8	263.25	183.90	101	-0.121
9	239.75	111.30	107	0.169
10	238.25	108.90	110	0.314
11	213.25	111.30	131	1.331
12	213.25	108.90	125	1.040
13	210.25	86.30	122	0.895
14	214.75	86.30	123	0.943
15	186.30	86.75	91	-0.605
16	188.70	86.75	81	-1.089
17	166.10	89.75	117	0.653
18	163.70	85.25	122	0.895
19	191.10	64.75	81	-1.089
20	191.10	60.25	83	-0.992
21	163.70	63.25	81	-1.089
22	166.10	63.25	92	-0.556
23	140.80	85.25	102	-0.073
24	145.60	85.25	109	0.266
25	140.80	55.25	113	0.460
26	143.20	53.75	108	0.218
27	104.40	87.75	67	-1.766
28	106.80	87.75	86	-0.847
29	104.40	60.25	137	1.621
30	104.40	58.75	119	0.750
31	88.70	75.75	59	-2.153
32	91.10	75.75	69	-1.669
33	88.70	47.25	87	-0.798
34	83.90	45.75	74	-1.427

Gemiddelde waarde  $E[\bar{q}_c] = 81.08$

Standaarddeviatie  $\sigma(\bar{q}_c) = 16.46 \text{ kgf/cm}^2$

Tabel VII. Berekende autocorrelatieparameters voor gemiddelde sondeerweerstand.

type	$\hat{a}_0$	$E[a_0]$	$\sigma(a_0)$	$\hat{D}$	$E[D]$	$\sigma(D)$
mc	0.9	0.73	0.146	8	12.6	5.0
ec	0.9	0.86	0.067	18.5	19.0	4.3

mc = mechanische mantelconus, ec = gladde elektrische conus

$\hat{a}_0, \hat{D}$  =  $a_0$  en D-waarden met hoogste likelihoodscores

$E[a_0], E[D]$  verwachtingswaarden, bepaald uit verdeling van likelihoodscores

$\sigma(a_0), \sigma(D)$  standaardafwijkingen, bepaald uit verdeling van likelihoodscores

bladnummer : - 28 -  
 ons kenmerk: C0266483/14  
 datum : 1987-07-28



Tabel VIII. Likelihoodscores voor  $a_0$  en D op basis van de waargenomen gemiddelde conusweerstand (mechanische mantelconus).

$a_0$ waarden →	0.40	0.50	0.60	0.70	0.80	0.90	1.00
D waarden [m]							
2.0	.00033	.00040	.00047	.00056	.00066	.00079	.00095
3.5	.00068	.00097	.00139	.00202	.00230	.00454	.00677
5.0	.00110	.00177	.00289	.00484	.00834	.01414	.01651
6.5	.00165	.00289	.00517	.00947	.01747	.02892	.01210
8.0	.00225	.00419	.00787	.01488	.02719	.03840	.00215
9.5	.00285	.00550	.01056	.01985	.03396	.03666	.00009
11.0	.00342	.00673	.01298	.02382	.03702	.02838	.00001
12.5	.00388	.00773	.01487	.02635	.03659	.01902	.00000
14.0	.00418	.00836	.01589	.02695	.03298	.01113	.00000
15.5	.00427	.00849	.01580	.02542	.02695	.00560	.00000
17.0	.00412	.00807	.01460	.02206	.01989	.00238	.00000
18.5	.00377	.00720	.01256	.01763	.01331	.00086	.00000
20.0	.00327	.00605	.01009	.01304	.00811	.00027	.00000
21.5	.00271	.00483	.00761	.00896	.00454	.00008	.00000
23.0	.00217	.00367	.00542	.00574	.00233	.00002	.00000
24.5	.00169	.00269	.00367	.00345	.00110	.00000	.00000
26.0	.00129	.00191	.00238	.00195	.00048	.00000	.00000

$E[a_0] = 0.733$ ,  $\sigma(a_0) = 0.146$

$E[D] = 12.60$  m,  $\sigma(D) = 5.0$  m

Correlatie tussen  $a_0$  en D: -0.44

Onderstreepte waarde is maximum likelihoodscore en deze behoort bij:

$\hat{a}_0 = 0.90$  en  $\hat{D} = 8.0$  m.



Tabel IX. Likelihoodscores voor  $a_0$  en D op basis van de waargenomen gemiddelde conusweerstand (gladde elektrische conus).

$a_0$ waarden →	0.40	0.50	0.60	0.70	0.80	0.90	1.00
D waarden [m]							
2.0	.00000	.00000	.00000	.00000	.00000	.00000	.00001
3.5	.00001	.00001	.00002	.00003	.00004	.00007	.00009
5.0	.00002	.00003	.00007	.00014	.00030	.00055	.00054
6.5	.00003	.00007	.00018	.00045	.00112	.00238	.00083
8.0	.00004	.00012	.00033	.00097	.00288	.00673	.00037
9.5	.00005	.00016	.00051	.00169	.00580	.01467	.00005
11.0	.00007	.00021	.00070	.00258	.00995	.02710	.00000
12.5	.00008	.00025	.00090	.00353	.01504	.04391	.00000
14.0	.00008	.00028	.00105	.00439	.02025	.06287	.00000
15.5	.00009	.00030	.00114	.00499	.02450	.07994	.00000
17.0	.00009	.00030	.00116	.00523	.02692	.09107	.00000
18.5	.00008	.00028	.00111	.00510	.02712	.09386	.00000
20.0	.00008	.00026	.00102	.00469	.02529	.08817	.00000
21.5	.00007	.00024	.00090	.00407	.02197	.07567	.00000
23.0	.00006	.00021	.00076	.00338	.01783	.05924	.00000
24.5	.00006	.00017	.00063	.00267	.01354	.04204	.00000
26.0	.00005	.00015	.00051	.00203	.00962	.02679	.00000

$E[a_0] = 0.864$ ,  $\sigma(a_0) = 0.067$

$E[D] = 17.98$  m,  $\sigma(D) = 4.3$  m

Correlatie tussen  $a_0$  en D: 0.04

Onderstreepte waarde is maximum likelihoodscore en deze behoort bij:

$\hat{a}_0 = 0.90$  en  $\hat{D} = 8.0$  m.

### 3.4 Conclusies

#### a. Ten aanzien van de berekeningen op gegenereerde velden:

De methode werkt goed, ook bij relatief weinig waarnemingen, als deze maar voldoende informatie bevatten (al er maar voldoende tweetallen zijn aan te wijzen die onderling nog sterk gecorreleerd zijn). Indien de sterkst gecorreleerde waarnemingen zwak gecorreleerd zijn (met een correlatiecoëfficiënt van pakweg beneden 0.35), dan is de verdeling van likelihoodscores sterk gespreid over de kleine fluctuatieschalen. De methode is dan wel in staat om aan te geven beneden welke grens de fluctuatieschaal ligt, maar kan geen uitsluitsel geven waar precies beneden die grens. De reden hiervoor is dat de meting zelf deze informatie niet bevat. Het





aantal meetpunten is wel van belang. Als de meting goede informatie over het afvalgedrag van de autocorrelatie bevat, dan zal de schatting van de fluctuatieschaal beter zijn naarmate meer meetpunten beschikbaar zijn. Bevat de meting intrinsiek geen informatie over het afvalgedrag, dan helpen ook veel waarnemingen niet.

b. Ten aanzien van de sonderingen:

In het beschouwde veld is de variatiecoëfficiënt van de gemiddelde sondeerweerstand in de draagkrachtige zandlaag ongeveer 0.20. dat betekent dat met kans van zeg 0.95 de gemiddelde sondeerweerstand tussen 0.6 en 1.4 keer de gemiddelde waarde ligt, dit overigens ondanks het vermoeden dat het hier om een homogeen terrein ging. Ongeveer 7 a 10 procent van de variantie komt voor rekening van reproduceerbaarheidsfouten (onnauwkeurigheid van het meetinstrumentarium en/of terreinvarianties op zeer kleine schaal) de rest moet worden toegeschreven aan terreinvariantie. Uit de berekende autocorrelatieparameter  $D$  kunnen we afleiden hoe groot de omgeving is waarvoor een sondering representatief gesteld kan worden en welke de mogelijke fout is die gemaakt wordt door een sondering representatief te stellen voor een hei-locatie ver weg van die sondering. In principe ligt daarmee de weg open voor een economische afweging van kosten en baten van meer of minder terreinonderzoek (Calle [13]).

4. TRENDANALYSE BIJ NORMAAL VERDEELDE STOCHASTISCHE VELDEN.

4.1 Inleiding.

Een belangrijke beperking voor het toepassen van de statistische berekeningen die in de voorgaande hoofdstukken zijn beschouwd, is de aanname dat we te maken hebben met een homogeen stochastisch veld. Een in de praktijk vaak voorkomende "inhomogeniteit" betreft de aanwezigheid van onderliggende gemiddelde trend in het variatiepatroon van een of andere terreingrootheid. Om de statistische theorie toch te kunnen toepassen is het nodig om deze gemiddelde trend te herkennen en uit het variatiepatroon, voor zover middels puntwaarnemingen bekend, te elimineren. De gevonden trend wordt als deterministische bijdrage aan het variatiepatroon opgevat en het overblijvende deel als stochastische variatie op die trend. Statistische berekeningen hebben alleen betrekking op deze stochastische variaties.

Het herkennen van gemiddelde trend in een variatiepatroon, op basis van een aantal puntwaarnemingen, is in feite geen gemakkelijke zaak. Er zijn in het algemeen geen op fysische overwegingen gestoelde criteria op grond waarvan een variatiepatroon van een fysische terreingrootheid kan worden opgedeeld in een gemiddelde trend en stochastische variaties. Dit betekent dat zo'n verdeling op puur mathematische grond zal moeten gebeuren en daarom altijd enigszins arbitrair zal zijn. Immers elk onderscheid tussen trend en stochastische variaties is, bij het ontbreken van een fysische

overweging, kunstmatig. Dit is overigens geen ramp, immers we moeten er alleen maar voor zorgen dat het variatiepatroon dat we overhouden na eliminatie van wat we als trend onderkennen, zodanig is dat dit beschreven kan worden als een homogeen stochastisch veld. Hierin mag niet meer een onderliggende gemiddelde trend zichtbaar zijn.

Een voor de hand liggende methode om gemiddelde trends op te sporen op basis van puntwaarnemingen is de kleinste kwadraten analyse. Daarbij worden de parameters van een aangenomen analytisch expressie voor de trendfunctie zodanig bepaald dat de door die expressie beschreven trendfunctie het best aansluit bij de waarnemingen van het variatiepatroon. Dit in die zin dat de som van de kwadraten van de afstanden tussen waarnemingen en trendfunctie minimaal is. De aldus gevonden trendfunctie wordt opgevat als gemiddelde trend. De trendbijdrage per waarneming kan worden berekend en van de waargenomen waarde worden afgetrokken. Dit levert de stochastische variatie per waarneming. Echter de op deze wijze gevonden waarden voor de parameters in de trendfunctie zijn sterk afhankelijk van de toevallige waarnemingen. Bijvoorbeeld, als de waarnemingen realisaties zijn van een op zich trendvrij homogeen stochastisch veld, zullen we toch een trendfunctie vinden die ongelijk nul is. Daarom moet getoetst worden in hoever de gevonden trendfunctie significant ongelijk nul is. Dit kan gebeuren door de spreiding van de trendparameters bij de gegeven verdeling van de waarnemingen over het veld, bij een trendvrij veld, te berekenen. Met behulp van de gevonden spreidingen (standaardafwijkingen) kunnen waarschijnlijkheidsgebieden voor de trendparameters bij een trendvrij veld aangeven (bijvoorbeeld 0.9 betrouwbaarheidsintervallen). Indien een op basis van de waarnemingen berekende trendparameter binnen het overeenkomstige waarschijnlijkheidsgebied ligt nemen we aan dat de gevonden parameterwaarde niet significant afwijkt van de waarde die hoort bij een trendvrij veld. De met die parameter overeenkomende bijdrage aan de trendfunctie wordt dan genegeerd. Indien deze situatie zich voordoet voor alle trendparameters, dan concluderen we dat het variatiepatroon geen significante trend vertoont.

Als eerste wordt in dit hoofdstuk de beschrijving van de methode gegeven die gebaseerd is op het klassieke concept van de kleinste kwadraten methode, waarbij als onderdelen van de trendfunctie de basispolynomen  $(1, x, y, x^2, y^2, xy)$  worden gekozen voor een tweedimensionaal veld. De berekeninggang voor een- of drie- of zelfs  $n$ -dimensionale velden is geheel analoog, evenals de berekeninggang bij de keuze van een basispolynomen met hogere graad. Deze beschrijving heeft het nadeel dat de trendparameters die gevonden worden onderling gecorreleerd zijn, hetgeen de beoordeling van de significantie van de parameterwaarden bemoeilijkt. Daarom is vervolgens overgestapt op een methode die gebaseerd is op orthonormale polynomen. De definitie van orthonormaliteit is die welke wordt voortgebracht door het kleinste kwadraten concept.

Een niet opgeloste moeilijkheid bij de analyse van significantie van de trend is de volgende. De variantie van het veld en de varianties van de trendparameters bij een trendvrij veld worden berekend onder aanname dat de waarnemingen onderling ongecorreleerd zijn. Bij de velden die we hier willen onderzoeken is hieraan strikt genomen niet

aan voldaan, immers de gehele exercitie wordt uitgevoerd ten behoeve van het bepalen van de onderlinge correlatie tussen waarnemingen en we moeten derhalve dus wel aannemen dat onderlinge correlatie aanwezig kan zijn. Naar verwachting is het effect van die correlatie gering indien het aantal waarnemingen dat ten grondslag ligt aan de analyse relatief groot is en er zich daaronder dus ook een relatief groot aantal ongecorreleerde waarnemingsparen bevindt. Het effect van het meenemen van onderlinge correlatie tussen de waarnemingen zal overigens leiden tot grotere varianties van de trendparameters en derhalve tot een grotere kans op het niet verwerpen van de hypothese dat een veld trendvrij is.

#### 4.2 Test voor de afwezigheid van significante trend.

Veronderstel dat we beschikken over een aantal waarnemingen van een of andere eigenschap in een continu variërend veld. Deze waarnemingen geven we aan met  $\{\delta_i\} = \{\delta(\underline{x}_i)\}$  ( $i=1..n$ ), waarbij de  $\underline{x}_i$  plaatsvectoren in het veld voorstellen. Verondersteld wordt dat de veldvariabele  $\delta(\underline{x})$  een grillig variërende grootte is, die kan worden opgevat als een realisatie van een stochastisch veld, dat bestaat uit een deterministische component, de zogenaamde trend of drift en een stochastische component. In dit hoofdstuk willen we nagaan hoe met behulp van waarnemingen bepaald kan worden of in het veld een significante trend aanwezig is en zo ja, hoe die dan kan worden gekwantificeerd.

Om na te gaan of er een significante trend aanwezig is veronderstellen we in eerste instantie dat het veld trendvrij is en controleren op basis van de waarnemingen of deze hypothese verworpen dient te worden. Onder de veronderstelling dat het veld trendvrij is kunnen we schattingen voor de verwachtingswaarde en de standaardafwijking van de stochast  $\delta$  bepalen:

$$\bar{\delta} = \frac{1}{n} \sum_{i=1}^n \delta_i \quad (4.1)$$

$$s = \left\{ \frac{1}{n-1} \sum_{i=1}^n (\delta_i - \bar{\delta})^2 \right\}^{1/2} \quad (4.2)$$

Het rekenkundig gemiddelde van de steekproef  $\{\delta_i\}$  is een schatter voor de verwachtingswaarde  $\mu(\delta)$  en de standaardafwijking  $s$  van de steekproef is een schatter voor de standaardafwijking  $\sigma(\delta)$  van het veld.

Als er wel een trend in het veld aanwezig zou zijn, dan zouden we die als volgt berekenen. We veronderstellen dan een formele uitdrukking voor de beschrijving van die trend in termen van een volledig polynoom van een of andere graad. Voor een tweedimensionaal veld ( $\underline{x} = (x,y)$ ) en een tweede graads polynoom is de formele uitdrukking voor de trendfunctie:

$$D(x,y) = a_1 + a_2 x + a_3 y + a_4 x^2 + a_5 y^2 + a_6 xy \quad (4.3)$$

De trendparameters  $a_1 \dots a_6$  worden met behulp van de methode van de kleinste kwadraten bepaald. Daarbij wordt gezocht naar die combinatie van  $a_1 \dots a_6$  waarbij de som van de kwadraten van de verschillen tussen de waargenomen waarden  $\{\delta_i\}$  en de "theoretische" waarden  $D(x_i, y_i)$  minimaal is, dus:

$$R(a_1, a_2 \dots a_6) = \sum_{i=1}^n (\delta_i - D(x_i, y_i))^2 \quad (4.4)$$

minimaliseer R onder variatie van  $a_1 \dots a_6$

De oplossing van dit minimaliseringsprobleem wordt gevonden door R naar de parameters  $a_1, a_2 \dots a_6$  te differentieren en deze afgeleiden gelijk nul te stellen. Dit levert een lineair algebraïsch stelsel op waaruit de trendparameters zijn op te lossen. Dit stelsel ziet er als volgt uit voor de uitdrukking voor D zoals in vergelijking (4.3) genoemd is:

$$C \hat{\underline{a}} = \underline{f} \quad (4.5)$$

waarin de (nxn) matrix:

$$C = \begin{pmatrix} n & \sum x_i & \sum y_i & \sum x_i^2 & \sum y_i^2 & \sum x_i y_i \\ & \sum x_i^2 & \sum x_i y_i & \sum x_i^3 & \sum x_i y_i^2 & \sum x_i^2 y_i \\ & & \sum y_i^2 & \sum x_i^2 y_i & \sum y_i^3 & \sum x_i y_i^2 \\ \text{symmetrisch} & & & \sum x_i^4 & \sum x_i^2 y_i^2 & \sum x_i^3 y_i \\ & & & & \sum y_i^4 & \sum x_i y_i^3 \\ & & & & & \sum x_i^2 y_i^2 \end{pmatrix} \quad (4.6)$$

en  $\hat{\underline{a}}$  de vector:

$$\hat{\underline{a}} = (\hat{a}_1, \hat{a}_2, \hat{a}_3, \hat{a}_4, \hat{a}_5, \hat{a}_6)^T \quad (4.7)$$

waarin de  $\hat{a}_i$  schattingen van de echte trendparameters  $a_i$  zijn. In (4.6) is verder de vector  $\underline{f}$ :

$$\underline{f} = (\sum \delta_i, \sum \delta_i x_i, \sum \delta_i y_i, \sum \delta_i x_i y_i, \sum \delta_i x_i^2, \sum \delta_i x_i y_i, \sum \delta_i y_i^2)^T \quad (4.8)$$

De T als superscript bij een vector of matrix betekent "getransponeerd", dat wil zeggen dat een "liggende" vector door toevoeging van dit superscript verandert in een "staande" vector v.v. en een matrix verandert door toevoeging van dit superscript in een andere matrix, waarvan de rijen gelijk zijn aan de kolommen van de originele matrix.

Merk op dat voor een symmetrische matrix geldt:  $C = C^T$ .

De vector  $\underline{f}$  kunnen we schrijven als:

$$\underline{f} = \sum \underline{v}_i \delta_i \quad (4.9)$$

waarin:

$$\underline{v}_i = (1, x_i, y_i, x_i^2, x_i y_i, y_i^2)^T \quad (4.10)$$

De oplossing van het stelsel (4.5) luidt, in formele zin:

$$\underline{\hat{a}} = C^{-1} \underline{f} = C^{-1} \sum \underline{v}_i \delta_i \quad (4.11)$$

waarin  $C^{-1}$  de inverse matrix van  $C$  is.

Indien, zoals voor onze test op de afwezigheid van significante trend verondersteld is, het veld trendvrij is, dus  $a_2=0, a_3=0 \dots a_6=0$ , dan zullen we voor de uitkomst van de berekening volgens (4.11) toch waarden van  $\hat{a}_2 \dots \hat{a}_6$  ongelijk nul vinden, omdat de meetwaarden  $\delta_i$  een toevallige waarde hebben (ze zijn immers getrokken uit een populatie met een bepaalde kansdichtheidsverdeling). Als de kansdichtheidsverdeling normaal is, dan zijn ook de  $\hat{a}_i$  normaal verdeelde stochasten. Deze hebben een verwachtingswaarde en een standaardafwijking. De relatie tussen deze verwachtingswaarden en standaardafwijkingen en de verwachtingswaarde  $\mu(\delta)$  en standaardafwijking  $\sigma(\delta)$  van het veld wordt als volgt bepaald. De meetwaarden kunnen we formeel als volgt schrijven:

$$\delta_i = \underline{v}_i^T \underline{a} + \omega_i \quad (4.12)$$

waarin  $\underline{a} = (a_1, a_2, a_3, a_4, a_5, a_6)^T$  de vector met de echte trendparameters is, en de  $\omega_i$  de stochasten die overblijven na eliminatie van de trend.

Deze hebben een verwachtingswaarde gelijk 0 en een standaardafwijking  $\sigma(\omega) = \sigma(\delta)$ . In ons geval is  $a_1 = \mu(\delta)$  en  $a_2 = a_3 = \dots a_6 = 0$ . Invullen van vergelijking (4.12) in levert, als we daarbij in beschouwing nemen dat de matrix  $C$  geschreven kan worden als:

$$C = \sum_i \underline{v}_i \underline{v}_i^T \quad (4.13)$$

het volgende op:

$$\begin{aligned} \underline{\hat{a}} &= C^{-1} \sum_i \underline{v}_i \underline{v}_i^T \underline{a} + C^{-1} \sum_i \underline{v}_i \omega_i \\ &= \underline{a} + C^{-1} \sum_i \underline{v}_i \omega_i \end{aligned} \quad (4.14)$$

De verwachtingswaarde van  $\underline{\hat{a}}$  is dus:

$$E[\hat{\underline{a}}] = \underline{a} \quad (4.15)$$

en de zogenaamde covariantiematrix:

$$\begin{aligned} \text{cov}(\hat{\underline{a}}) &= E[(\hat{\underline{a}} - E[\hat{\underline{a}}])(\hat{\underline{a}} - E[\hat{\underline{a}}])^T] = E[(C^{-1} \sum_i \underline{v}_i \omega_i)(C^{-1} \sum_j \underline{v}_j \omega_j)^T] \\ &= C^{-1} \left( \sum_i \sum_j \underline{v}_i E[\omega_i \omega_j] \underline{v}_j^T \right) C^{-1T} \end{aligned} \quad (4.16)$$

Indien we veronderstellen dat de waarnemingen onderling onafhankelijk zijn, dat wil zeggen  $E[\omega_i \omega_j] = 0$  als  $i$  ongelijk aan  $j$  is en  $E[\omega_i \omega_i] = \sigma^2(\omega)$  ( $= \sigma^2(\delta)$ ), dan vinden we:

$$\text{cov}(\hat{\underline{a}}) = C^{-1} \sigma^2(\omega) \quad (4.17)$$

In de covariantiematrix staan de volgende getallen:

$$\text{cov}(\hat{\underline{a}}) = \begin{vmatrix} \sigma^2(\hat{a}_1) & \text{cov}(\hat{a}_1, \hat{a}_2) & \dots & \text{cov}(\hat{a}_1, \hat{a}_s) \\ & \sigma^2(\hat{a}_2) & \dots & \text{cov}(\hat{a}_2, \hat{a}_s) \\ & & \ddots & \\ \text{symmetrisch} & & & \sigma^2(\hat{a}_s) \end{vmatrix} \quad (4.18)$$

waarin  $\text{cov}(ab)$  de covariantie van de stochasten  $a$  en  $b$  voorstelt en die is als volgt gedefinieerd:

$$\text{cov}(ab) = E[(a - E[a])(b - E[b])] \quad (4.19)$$

Indien we moeten veronderstellen dat de waarnemingen niet (alle) onderling onafhankelijk zijn, dan moet in vergelijking (4.16) gerekend worden met:

$$E[\omega_i \omega_j] = \sigma^2(\omega) \rho_{ij}$$

waarin  $\rho_{ij}$  de correlatie tussen de waarnemingen is. Die correlatie is niet a priori bekend, immers de trendanalyse wordt uitgevoerd om eventuele trend te verwijderen voordat de analyse voor de berekening van de autocorrelatie wordt uitgevoerd. Het meenemen van eventuele correlatie tussen de waarnemingen zal wel tot gevolg hebben dat de standaardafwijkingen van de polynoomcoëfficiënten groter worden ten opzichte van de situatie waarin die correlatie niet wordt meegenomen. Dit betekent dat dan eerder de trend als niet significant wordt beoordeeld. Met andere woorden, rekenen we de standaardafwijkingen van de polynoomcoëfficiënten uit onder de aanname van onafhankelijkheid van de waarnemingen en vinden we op grond daarvan dat er geen significante trend is, dan zal die uitspraak niet veranderen als wel correlatie zou zijn meegenomen. Is de uitkomst van de analyse dat er wel significante trend is, dan kan wellicht het tegengestelde worden gevonden als wel correlatie wordt meegenomen.

Recapitulerend hebben we het volgende gevonden. Als we een trendberekening uitvoeren op een verzameling waarnemingen van een trendvrij veld (m.a.w.  $a_1 = \mu(\delta)$  en  $a_2, \dots, a_6 = 0$ ), dat homogeen normaal verdeeld is, dan vinden we, omdat de  $\delta_i$  stochasten zijn toch waarden ongelijk nul

voor de parameters  $\hat{a}_1, \dots, \hat{a}_5$ . Deze parameterwaarden zijn normaal verdeeld, met verwachtingswaarde gelijk nul (formule 4.15) en variantie ongelijk nul (formule 4.17). We kunnen voor elke trendparameter nagaan hoe groot de kans is dat de absolute waarde hiervan kleiner is dan een bepaalde gekozen waarde. Bijvoorbeeld de kans dat  $|\hat{a}_1|$  kleiner is dan  $1.96\sigma(\hat{a}_1)$  is 0.95. Dit volgt direct uit

de verdelingsfunctie van  $\hat{a}_1$ . We kunnen nu de zaak ook omkeren: de kans dat, onder de aanname  $E[\hat{a}_1] = 0$  ( $i=2 \dots 6$ ) een  $\hat{a}_i$  waarde uit de regressie analyse wordt gevonden die groter is dan  $1.96\sigma(\hat{a}_i)$  is gelijk aan 0.05.

Als dus toch zo'n  $\hat{a}_i$  waarde wordt gevonden dan was dus met grote kans de veronderstelling  $E[\hat{a}_i] = 0$  onjuist. Deze redenering is de basis voor de test op het trendvrij zijn van het veld, die als volgt is geformuleerd:

- De nulhypothese is dat  $a_i = 0$  ( $i=2 \dots 6$ ).
- Deze hypothese toetsen we tegen de alternatieve hypothese, die luidt dat  $a_i \neq 0$ .
- Toetscriterium: De nulhypothese wordt niet verworpen als  $|\hat{a}_i|$  ( $i=2 \dots 6$ ) kleiner is dan de bijbehorende  $1.96\sigma(\hat{a}_i)$ . In het andere geval wordt de nulhypothese verworpen en daarmee dus de alternatieve hypothese niet verworpen.

De kans dat de nulhypothese niet wordt verworpen als het veld inderdaad trendvrij is is 0.95 en de kans dat de nulhypothese toch wordt verworpen onder die omstandigheden (dit heet een fout van de eerste soort) is 0.05. Deze kans wordt het significantieniveau of de onbetrouwbaarheidsdrempel van de test genoemd. Het niet verwerpen van de nulhypothese terwijl er toch echte trend is wordt een fout van de tweede soort genoemd. De kans dat de nulhypothese wordt verworpen terwijl er echte trend is wordt het onderscheidingsvermogen van de test genoemd. Voor een goede statistische toets willen we:

- een lage onbetrouwbaarheidsdrempel
- een hoog onderscheidingsvermogen

dat wil zeggen, een kleine kans dat de nulhypothese ten onrechte wordt verworpen en een grote kans dat alternatieve hypothese terecht niet wordt verworpen. Deze twee kansen geven de scherpte van de test aan. De twee scherpte-eisen zijn vaak tegenstrijding en er moet dan ook een

compromis worden gezocht. Zoals zal blijken is dat ook bij onderhavige test het geval.

We beschouwen de fout van de tweede soort en gaan na hoe groot de kans hierop is. Veronderstel dat echte trend aanwezig is en er dus een  $E[\hat{a}_i] = a_i \neq 0$  is. De nulhypothese wordt niet verworpen indien  $|\hat{a}_i| < 1.96\sigma(\hat{a}_i)$ . De kans hierop is:

$$P(\text{fout } 2^{\text{e}} \text{ soort}) = \Phi\left[-\frac{a_i}{\sigma(\hat{a}_i)} + 1.96\right] - \Phi\left[-\frac{a_i}{\sigma(\hat{a}_i)} - 1.96\right] \quad (4.20)$$

waarin  $\Phi[\ ]$  de standaardnormale kansverdelingsfunctie is. Als bijvoorbeeld  $a_i = 1.96\sigma(\hat{a}_i)$ , dan is de kans op een fout van de tweede soort bijna 0.5. In het algemeen als  $a_i = t \sigma(\hat{a}_i)$  dan is de kans op een fout van de tweede soort:

$$P(\text{fout } 2^{\text{e}} \text{ soort}) = \Phi[t+1.96] - \Phi[t-1.96] \quad (4.23)$$

Het onderscheidingsvermogen is 1.0 minus deze kans. Het onderscheidingsvermogen varieert dus tussen 0.05 en 1.0, afhankelijk van de waarde van  $t$ . Voor kleine waarden van  $t$  is er dus sprake van slecht onderscheidingsvermogen. We zoeken daarom naar een andere maat om de scherpte van de test aan te geven. We berekenen de kans dat de nulhypothese niet wordt verworpen als de absolute waarde van de te onderzoeken trendcoëfficiënt groter is dan  $t$  keer de betreffende standaardafwijking. Deze kans is gelijk aan:

$$P(\text{fout } 2^{\text{e}} \text{ soort} \mid |a_i| > t\sigma(\hat{a}_i)) = \frac{\int_t^\infty \phi(u+b) - \phi(u-b) du}{\int_0^\infty \phi(u+b) - \phi(u-b) du}$$

$$= \left(\frac{t}{b}+1\right) \phi(t+b) - \left(\frac{t}{b}-1\right) \phi(t-b) - 1 + \frac{e^{-1/2(t+b)^2} - e^{-1/2(t-b)^2}}{b\sqrt{2\pi}}$$

... (4.22)

met  $b=1.96$ . Kiezen we  $t=1.96$ , dan vinden we een kans op een fout van de tweede soort van 0.20, dat wil zeggen dat indien de werkelijke trendcoëfficiënt in absolute waarde groter is dan 1.96 keer z'n overeenkomstige standaardafwijking, er een kans bestaat van 0.20 dat de nulhypothese (geen trend!) toch niet wordt verworpen. Kiezen we  $t=3$ , dan neemt die kans af tot 0.12. Omgekeerd geldt dat als de nulhypothese niet wordt verworpen, dat dan de echte waarde  $a_i$  van de trendcoëfficiënt met kans 0.90 kleiner is dan  $3\sigma(\hat{a}_i)$ . Het



onderscheidingsvermogen voor die waarden van de trendcoefficient is 0.90.

#### 4.3 Toepassing.

De regressie analyse van de vorige paragraaf zullen we toepassen op de mechanische mantelconus sonderingen die in hoofdstuk 3 beschouwd zijn. In principe kunnen verschillende berekeningen worden gemaakt. In eerste instantie is een serie berekeningen gemaakt waarbij we het aantal mee te nemen termen in de regressievergelijking steeds laten toenemen. Dus bij de eerste berekening alleen de constante term (1) en  $x$ , bij de tweede berekening  $1, x$  en  $y$ , bij de derde  $1, x, y$  en  $x^2$ , enzovoorts. De berekeningsresultaten zijn weergegeven in tabel X. In deze tabel zijn aangegeven het berekeningsgeval (de polynoomtermen die zijn meegenomen in de analyse), de berekende polynoomcoefficienten en de standaardafwijkingen ervan. In de meest rechtse kolom is tevens de residuele standaardafwijking gegeven, dat is de standaardafwijking van de waarnemingen ten opzichte van het berekende regressievlak. Door de gevonden waarden van de polynoomcoefficienten te vergelijken met de standaardafwijkingen ervan bepalen we of de betreffende polynoomterm significant is of niet (significant indien de absolute waarde van de polynoomterm groter dan 1.96 keer  $z$ 'n standaardafwijking). We zien de volgende resultaten. Bij de eerste berekening (met termen 1 en  $x$ ) blijkt  $a_2$  niet significant te zijn. Bij de tweede berekening ( $1, x, y$ ) blijken zowel  $a_2$  als  $a_3$  niet significant en bij de derde berekening ( $1, x, y, x^2$ ) zijn ook alle polynoomcoefficienten niet significant. Hetzelfde vinden we ook bij de volgende berekening ( $1, x, y, x^2, y^2$ ). Echter, nemen we alle polynoomtermen mee, dan blijkt ineens dat  $a_4$  net wel/net niet en  $a_5$  wel significant te zijn (in tegenstelling tot de vorige berekening), evenals de nieuw toegevoegde  $a_6$ . Het al dan niet significant zijn van een polynoomcoefficient hangt dus kennelijk af van welke polynoomtermen in de regressieberekening zijn ingevoerd. We kunnen ook regressie berekeningen uitvoeren waarbij de verschillende polynoomtermen stuk voor stuk als enige in de berekening worden meegenomen. De resultaten zijn eveneens genoemd in tabel X. Uit deze berekeningen zou blijken dat geen van de termen significant is. We concluderen dat een regressie analyse op de voorgestelde wijze niet leidt tot eenduidige uitkomsten en daarom weinig houvast biedt. Om die reden is ook overgegaan tot de procedure die in de volgende paragraaf wordt afgeleid.

Tabel X. Uitkomst van de regressieberekeningen volgens de procedure van paragraaf 4.1.

polynoom coefficienten<sup>1)</sup>, standaardafwijkingen<sup>2)</sup> (tussen haakjes) en residuele standaardafwijkingen<sup>3)</sup>  $\sigma_{res}$ :

reken- geval:	$\hat{a}_0$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$	$\hat{a}_5$	$\hat{a}_6$	$\sigma_{res}$
1 <sup>4)</sup>	73.78	0.039 (0.042)					16.26
2	73.76	0.053 (0.073)	-0.025 (0.109)				16.25
3	68.06	0.113 (0.323)	-0.008 (0.140)	-0.0002 (0.0010)			16.24
4	52.40	-0.078 (0.350)	0.626 (0.460)	0.0003 (0.0010)	-0.0027 (0.0019)		15.74
5	57.61	0.201 (0.363)	0.157 (0.492)	-0.0040 (0.0020)	-0.0140 (0.0040)	0.0150 (0.0040)	13.43
6	81.71			-0.0040 (0.0010)	-0.0130 (0.0040)	0.0150 (0.0040)	13.58
7	77.15		0.039 (0.063)				16.37
8	77.24			0.0001 (0.0001)			16.28
9	79.93				0.0001 (0.0003)		16.43
10	78.39					0.0001 (0.0002)	16.34

1) berekend op basis van de waarnemingen

2) bij veronderstelde afwezigheid van trend

3) standaardafwijking van meting ten opzichte van berekende trend

4) geval 1: basis polynomen 1, x

2	1, x, y
3	1, x, y, x <sup>2</sup>
4	1, x, y, x <sup>2</sup> , y <sup>2</sup>
5	1, x, y, x <sup>2</sup> , y <sup>2</sup> , xy
6	1, x <sup>2</sup> , y <sup>2</sup> , xy
7	1, y
8	1, x <sup>2</sup>
9	1, y <sup>2</sup>
10	1, xy

#### 4.4 Trendanalyse gebaseerd op orthonormale polynomen.

De regressieanalyse, zoals behandeld in paragraaf 4.2, heeft een belangrijk nadeel. Stelt men zich voor dat een eerste analyse wordt uitgevoerd met alleen het eerstegraads polynoom, dus (in twee dimensies):

$$p(x,y) = a_1 + a_2 x + a_3 y \quad (4.23)$$

dan levert deze analyse de uitkomsten  $\hat{a}_1^1$ ,  $\hat{a}_2^1$  en  $\hat{a}_3^1$  op. Wordt vervolgens een analyse uitgevoerd met een volledig tweedegraads polynoom:

$$p(x,y) = a_1 + a_2 x + a_3 y + a_4 x^2 + a_5 y^2 + a_6 xy \quad (4.24)$$

dan levert dit uitkomsten  $\hat{a}_1^q$ ,  $\hat{a}_2^q, \dots, \hat{a}_6^q$  op. Merk op dat de uitkomsten van de eerste analyse zijn aangegeven met superscript 1 en de uitkomsten van de tweede analyse met q. De reden hiervoor is dat in het algemeen de  $\hat{a}_i^1 \neq \hat{a}_i^q$ , zoals al blijkt uit het rekenvoorbeeld van paragraaf 4.3. Dit is een belangrijk nadeel van de methode, die z'n oorzaak vindt in het feit dat de basispolynomen  $\{1, x, y, x^2, y^2, xy\}$  een niet orthonormale basis vormen in de polynoomruimte en dat de uitkomsten van de kleinste kwadraten analyses geen orthogonale projecties zijn (in de zin van de kleinste kwadraten methode). Dit euvel willen we verhelpen door, uitgaande van de basispolynomen een basis van orthonormale polynomen te construeren en de trendanalyse te betrekken op deze polynomen.

Uitgangspunt is het stelsel basispolynomen dat boven genoemd is. De ruimte van polynomen kunnen we representeren door een vectorruimte  $\underline{a} = (a_1, a_2, \dots, a_6)^T$ , de coëfficiënten van een polynoom. Het bijbehorende polynoom kunnen we schrijven als:

$$p(x,y;\underline{a}) \stackrel{\text{notatie}}{=} p(\underline{a}) = \underline{a}^T \underline{v} \quad (4.25)$$

waarin  $\underline{v} = (1, x, y, x^2, y^2, xy)^T$  een formele vectornotatie is voor de basispolynomen. Als inwendig product in de ruimte van polynomen definiëren we de bilineaire functionaal:

$$[p(\underline{a}), p(\underline{b})] = \sum_{i=1}^n \underline{a}_i^T \underline{v}_i \underline{v}_i^T \underline{b} = \underline{a}^T (\sum_{i=1}^n \underline{v}_i \underline{v}_i^T) \underline{b} \quad (4.26)$$

waarin de  $\underline{v}_i$  de vectorwaarden zijn van  $\underline{v}$  in de waarnemingspunten van de variabele die onderwerp van analyse is. Merk op dat de matrix  $\sum_{i=1}^n \underline{v}_i \underline{v}_i^T$  precies de (symmetrische) C matrix is die geïntroduceerd is in paragraaf 4.1. Deze matrix is, mits inverteerbaar, positief definit. Indien de matrix inverteerbaar is dan is het regressieprobleem goed gesteld en vice versa. Omdat C positief

definit is, is een eenduidige decompositie mogelijk (bijvoorbeeld met behulp van het Choleski algoritme):

$$C = G^T G \quad (4.27)$$

waarin G een bovendiagonaal matrix is.

We willen nu een in de in van bovengenoemd inproduct een orthonormale basis in de polynoomruimte definiëren, met andere woorden we zoeken naar een basis  $\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_6$ , waarvoor geldt dat:

$$[\tilde{e}_i, \tilde{e}_j] = 0 \text{ als } i \neq j \text{ en } [\tilde{e}_i, \tilde{e}_i] = 1 \quad (4.28)$$

Met behulp van de definitie van het inproduct voor de polynoomruimte en de decompositie kunnen we schrijven:

$$[\tilde{e}_i, \tilde{e}_j] = \tilde{e}_i^T G^T G \tilde{e}_j \quad (4.29)$$

waarin de vectoren  $\tilde{e}$  de polynoomcoëfficiënten van de orthonormale polynomen bevatten:

$$\tilde{e}_i = \tilde{e}_i^T \underline{v} \quad (4.30)$$

We kunnen deze vectoren representeren door de matrix  $\tilde{E}$ , waarvan de kolommen de kentallen van de vectoren bevatten. Als eis voor een orthonormale basis vinden we dan:

$$\tilde{E}^T G^T G \tilde{E} = I \quad (4.31)$$

waarin I de (in dit geval 6x6) eenheidsmatrix is. Aan deze voorwaarde wordt onder andere voldaan indien:

$$\tilde{E} = G^{-1} \quad (4.32)$$

De matrix G is een bovendiagonaal matrix en derhalve is ook de matrix E een bovendiagonaal matrix. De orthonormale basispolynomen die we zoeken zijn derhalve:

$$\begin{aligned} \tilde{e}_1 &= \tilde{E}_{11} \\ \tilde{e}_2 &= \tilde{E}_{12} + \tilde{E}_{22} x \\ \tilde{e}_3 &= \tilde{E}_{13} + \tilde{E}_{23} x + \tilde{E}_{33} y \\ \tilde{e}_4 &= \tilde{E}_{14} + \tilde{E}_{24} x + \tilde{E}_{34} y + \tilde{E}_{44} x^2 \end{aligned} \quad (4.33)$$

bladnummer : - 42 -  
 ons kenmerk: C0266483/  
 datum : 1987-05-22



$$\tilde{e}_5 = \tilde{E}_{15} + \tilde{E}_{25} x + \tilde{E}_{35} y + \tilde{E}_{45} x^2 + \tilde{E}_{55} y^2$$

$$\tilde{e}_6 = \tilde{E}_{16} + \tilde{E}_{26} x + \tilde{E}_{36} y + \tilde{E}_{46} x^2 + \tilde{E}_{56} y^2 + \tilde{E}_{66} xy$$

In formele matrix vector notatie kunnen we deze vergelijkingen schrijven als:

$$\underline{\tilde{e}} = \tilde{E}^T \underline{v} \quad (4.34)$$

Waarin  $\underline{\tilde{e}}$  de formele vectornotatie is voor de orthonormale polynomen.

Een willekeurig polynoom in de polynoomruimte (tot de macht 2) wordt gerepresenteerd door:

$$p(x,y) = b_1 \tilde{e}_1 + b_2 \tilde{e}_2 + \dots + b_6 \tilde{e}_6 \quad (4.35)$$

Als in de punten  $(x_i, y_i)$  de waarde  $\omega_i$  van de veldvariabele is gemeten, dan kan op eenvoudige wijze de, in de zin van de kleinste kwadraten methode, de best passende set van polynoomcoëfficiënten  $\underline{b} = (b_1, \dots, b_6)^T$  worden bepaald, namelijk:

$$b_k = (\tilde{e}_k, \omega) = \sum_{i=1}^n \tilde{e}_k(x_i, y_i) \omega_i \quad (k = 1 \dots 6) \quad (4.36)$$

of in matrix-vector notatie:

$$\underline{b} = \sum_{i=1}^n \tilde{E}^T \underline{v}_i \omega_i = \tilde{E}^T \sum_{i=1}^n \underline{v}_i \omega_i \quad (4.37)$$

Het verband tussen de vector  $\underline{a}$ , werkend met de basispolynomen en de vector  $\underline{b}$ , werkend met de orthonormale polynomen, volgt uit beider polynoomrepresentaties en is:

$$\underline{a} = \tilde{E} \underline{b} = G^{-1} \underline{b} \quad (4.38)$$

dus:

$$\underline{a} = G^{-1} (G^{-1T} \sum_{i=1}^n \underline{v}_i \omega_i) = (G^T G)^{-1} \sum_{i=1}^n \underline{v}_i \omega_i = C^{-1} \sum_{i=1}^n \underline{v}_i \omega_i \quad \dots (4.39)$$

en de laatste term is precies de uitdrukking die we gevonden hebben bij de afleiding in paragraaf 4.2. Hieruit blijkt dat de in deze paragraaf voorgestelde rekenprocedure voor het bepalen van de trend feitelijk het kleinste kwadraten algoritme is. Het verschil is dat de trendcoëfficiënten niet afhankelijk zijn van het aantal van de in beschouwing genomen basispolynomen is, hetgeen bij de analyse in paragraaf 4.1 wel het geval was.

De standaardafwijkingen van de trendcoëfficiënten  $b_i$  ( $i=1\dots 6$ ), uitgaande van een trendvrij veld (nulhypothese), vinden we als volgt:

$$\begin{aligned} E[\underline{b} \underline{b}^T] &= G^{-1T} \left( \sum_{i=1}^n v_i \omega_i \right) \left( \sum_{j=1}^n \omega_j v_j^T \right) G^{-1} \\ &= G^{-1T} \left( \sum_{i=1}^n \sum_{j=1}^n v_i E[\omega_i \omega_j] v_j^T \right) G^{-1} \end{aligned} \quad (4.40)$$

Er geldt:

$$E[\omega_i \omega_j] = \sigma^2(\omega) \rho_{ij} \quad (4.41)$$

waarin  $\sigma^2(\omega)$  de variantie is en  $\rho_{ij}$  de correlatie tussen de waarnemingen. Indien we veronderstellen dat de waarnemingen onderling niet of slechts zwak gecorreleerd zijn dan volgt:

$$\begin{aligned} E[\underline{b} \underline{b}^T] &= \sigma^2(\omega) G^{-1T} \left( \sum_{i=1}^n v_i v_i^T \right) G^{-1} = \sigma^2(\omega) G^{-1T} G^T G G^{-1} = \sigma^2(\omega) I \\ &\dots (4.42) \end{aligned}$$

Dat wil zeggen dat de varianties van alle trendcoëfficiënten  $b_k$  ( $k=1\dots 6$ ) gelijk zijn aan de variantie van het veld en tevens dat de trendcoëfficiënten onderling ongecorrleerd zijn. We kunnen dus bij het toetsen op de significantie van de trendcomponenten direct toetsen op de standaardafwijking van het veld. In de trendanalyses die worden uitgevoerd wordt een trendcoëfficiënt significant ongelijk nul verondersteld als de berekende waarde 1.96 keer de standaardafwijking van het veld overschrijdt. De standaardafwijking wordt daarbij berekend onder de aanname dat er geen trend aanwezig is.

#### 4.5 Toepassing trendanalyse met behulp van orthonormale polynomen.

Als voorbeeld voor toepassing is weer het veld met de mechanische mantelconus sonderingen gekozen. De gevonden orthonormale polynomen (volgens vergelijking 4.33) luiden:

$$\tilde{e}_1 = 0.16667$$

$$\tilde{e}_2 = -0.48089 + 0.00257 x$$

$$\tilde{e}_3 = 0.00314 - 0.00358 x + 0.00661 y$$

$$\tilde{e}_4 = 1.80145 - 0.01914 x - 0.00565 y + 0.00006 x^2$$

$$\tilde{e}_5 = 0.66567 + 0.00813 x - 0.02690 y - 0.00002 x^2 + 0.00011 y^2$$

$$e_6 = 0.10712 + 0.00574 x - 0.00963 y - 0.00010 x^2 - 0.00023 y^2 + 0.00030 xy$$

en de gevonden polynoomcoefficienten zijn (zie vergelijking 4.36):

$$b_1 = 486.5$$

$$b_2 = 15.18$$

$$b_3 = -3.84$$

$$b_4 = -3.15$$

$$b_5 = -23.55$$

$$b_6 = 48.67$$

De standaardafwijking van het veld, bij veronderstelde afwezigheid van trend, is 16.46. De standaardafwijkingen van de polynoomcoefficienten is tenminste hieraan gelijk (is gelijk indien de waarnemingslocaties onderling ongecorrleerd zijn). Kiezen we als criterium voor het al dan niet verwerpen van een trendcomponent 1.96 keer de standaardafwijking, dan blijkt dat, afgezien van de constante term, alleen de laatste polynoomterm zo groot te zijn dat insignificantie van deze term verworpen moet worden. Laten we derhalve alle polynomen behalve de laatste (en natuurlijk de eerste constante term) weg uit de vergelijking voor het regressievlak, dan vinden we als uiteindelijk resultaat voor de regressievergelijking:

$$q(x,y) = 86.29 + 0.27969 x - 0.46887 y - 0.00472 x^2 - 0.01122 y^2 + 0.01466 xy$$

en de residuele standaardafwijking van de meting ten opzichte van dit vlak is  $\sigma_{res} = 14.26$ .

#### 4.6 Discussie.

De hier voorgestelde trendanalyse is in een bepaald opzicht onbevredigend. De regressiecoefficienten worden berekend onder de aanname dat de waarnemingen, of althans de overblijvende residuen ervan, onderling ongecorrleerd zijn. Dit is, in dit stadium van de verwerking van de meetresultaten noodzakelijk, omdat de onderlinge correlaties onbekend zijn. Het is redelijk om te veronderstellen dat die correlaties niet (alle) gelijk nul zijn, immers de gehele exercitie is erop gericht juist die te bepalen. Wanneer dat dan gebeurd is, dan zou eigenlijk de trendanalyse moeten worden overgedaan, nu rekening houdend met de correlaties, hetgeen zou kunnen resulteren in een andere trend, met als gevolg, bij de berekening van de correlaties op basis van de nieuwe residuen een andere autocorrelatiefunctie, enzovoort. Men zou natuurlijk kunnen besluiten een iteratief zoekproces voor trend en autocorrelatiefunctie op te

bladnummer : - 45 -  
ons kenmerk: C0266483/  
datum : 1987-05-22



GRONDMECHANICA  
DELFT

zetten, maar dit is erg onaantrekkelijk vanuit het oogpunt van bewerkelijkheid en rekenintensiteit (de berekening van de autocorrelatiefunctie op basis van residuen van waarnemingen is rekenintensief). Een andere mogelijkheid is om trendberekening en berekeningen van de autocorrelatiefunctie te integreren. In eerste instantie leek ook die mogelijkheid niet zo aantrekkelijk, het zou immers betekenen dat in de Bayesiaanse procedure het aantal te berekenen parameters zou moeten worden uitgebreid met het aantal trendcoëfficiënten. Bij nader inzien ligt een veel elegantere mogelijkheid voor de hand. Wanneer namelijk, zoals in de Bayesiaanse procedure gebeurt, door de keuze van de autocorrelatieparameter(s) de veronderstelde autocorrelatie tussen de residuen van de waarnemingen vastligt, dan bepaalt deze tegelijkertijd ook de beste schatting voor de trendfunctie en voor de variantie van de residuen. De relaties tussen beide en de autocorrelatiefunctie zijn eenduidig. Het is daarom mogelijk de trendanalyse te verenigen met de Bayesiaanse schattingsprocedure voor de autocorrelatieparameter(s), zonder in die procedure de trendcoëfficiënten apart mee te nemen. Dit opent interessante perspectieven, die echter nog nader moeten worden uitgewerkt. Zolang dit pad nog niet voldoende onderzocht is, is de beschreven procedure in dit hoofdstuk de meest aangewezen.

## 5. SAMENVATTING.

Voor een statistische beschrijving van variatiepatronen van fysische grootheden in een terrein kan gebruik gemaakt worden van het concept van homogene stochastische velden. Daarbij wordt het variatiepatroon van de fysische grootheid, na eliminatie van eventuele gemiddelde trend, opgevat als een realisatie van een homogeen stochastisch veld. Voorwaarde hiervoor is dat het variatiepatroon zodanig is dat ogenschijnlijk de waarden van de veldvariabele op elke plaats in het veld op een willekeurige manier tot stand zijn gekomen, zij het dat het variatiepatroon continu of nagenoeg continu kan zijn en dat de statistische verdeling van de variaties en de sterkte van de fluctuaties overal in het veld globaal dezelfde is. De plaatsonafhankelijkheid van de statistische verdeling van de variaties wordt in het model van het stochastisch veld gereflecteerd door de eigenschap dat de kansdichtheidsverdeling van de variabele op elke positie in het veld dezelfde is en de plaatsonafhankelijkheid van de sterkte van de fluctuaties wordt in het model gereflecteerd door de aanname dat correlaties tussen de stochastische veldvariabelen in elk tweetal punten in het veld uitsluitend afhangen van de afstanden (of afstandscomponenten in de verschillende veldrichtingen) tussen die twee punten in het veld.

De parameters van de kansdichtheidsverdeling van het veld en de vorm van de autocorrelatiefunctie moeten worden bepaald uit een steekproef van de waarden in het veld. Voor wat betreft de parameters van de kansdichtheidsverdeling, namelijk de verwachtingswaarde en de standaardafwijking, aannemende dat de kansdichtheidsverdeling Gaussisch is, kunnen schattingen gevonden worden met behulp van betrekkelijk eenvoudige rekenkundige bewerkingen op de waarnemingen. In dit rapport valt echter de nadruk op rekenkundige



bladnummer : - 46 -  
ons kenmerk: C0266483/  
datum : 1987-05-22



**GRONDMECHANICA  
DELFT**

methodieken om de autocorrelatiefunctie van het veld te schatten aan de hand van de waarnemingen. Het blijkt dat bestaande eenvoudige rekenkundige technieken om op basis van waarnemingsmateriaal de autocorrelatiefunctie te schatten falen, omdat de voor deze technieken benodigde aantallen waarnemingen om tot redelijk nauwkeurige schattingen te komen zeer groot zijn, in ieder geval veel groter dan het aantal waarnemingen dat redelijkerwijs bij een praktisch opgezet grondonderzoek beschikbaar zal zijn. Om dit bezwaar te omzeilen is een techniek ontwikkeld, die welliswaar betrekkelijk rekenintensief is, maar waarvoor het benodigde aantal waarnemingen betrekkelijk gering is. Deze methodiek is uiteengezet in hoofdstuk 3. De resultaten van experimenten met deze techniek, toegepast op zowel gegenereerde een en twee dimensionale velden als op velden die aan de praktijk zijn ontleend, met name sondeerweerstand, zijn bevredigend.

Variatiepatronen van fysische grootheden, zoals die in de praktijk worden aangetroffen vertonen naast de willekeurige (lees: stochastische) variaties veelal een gemiddelde trend. Deze trend moet allereerst opgespoord worden en vervolgens, indien significant, uit het waarnemingsmateriaal worden geelimineerd. In hoofdstuk 4 is een methode behandeld volgens welke die trendanalyse kan gebeuren en wordt ook een techniek aangedragen waarmee de significantie van een in de waarnemingen aanwezige trend kan worden beoordeeld.

6. REFERENTIES.

- [1] Calle, E.O.F. & W.J. Heijnen, 1984, Probabilistische stabiliteits-analyse (model en filosofie), LGM rapport CO-251433/6, i.o.v. Centrum Onderzoek Waterkeringen.
- [2] Calle, E.O.F., 1982, De keuze van meetnetfijnheid bij het in kaart brengen van grond of bodemeigenschappen, LGM rapport CO-256501/2 i.o.v. de Combinatie Speurwerk Baggertechniek.
- [3] Calle, E.O.F., 1986, A Geostatistical Approach to the Optimization of pre-dredging Site Survey, Proc. XI World Dredging Conference, Brighton.
- [4] Dixon, W.J. & F.J. Massey, 1962, Introduction to Statistical Analysis, Mc Graw-Hill, New York.
- [5] Vanmarcke, E., 1979, On the Scale of Fluctuation of Random Functions, MIT resrch. rep. R79-19.
- [6] Papoulis, A., 1965, Probability, Random Variables and Stochastic Processes, Mc Graw-Hill Kogakusha, Tokyo.
- [7] Klaver, E., B. Kuiper & A. Vrouwenvelder, 1978, Random Vibrations, T.U. Delft.
- [8] Tabba, M.M. & R.N. Yong, 1981, Mapping and Predicting of Soil Properties, Jrnl. Eng. Mech. Div. ASCE, EM5.
- [9] Journel, A.G. & Ch.I. Huijbregts, 1978, Mining Geostatistics, Academic Press, London.
- [10] Davies, J.C., 1973, Statistics and Data Analysis in Geology, John Wiley & Sons, New York.
- [11] Vrouwenvelder, A., 1985, Bayesiaanse schatting van de fluctuatieschaal van een normaal verdeeld stochastisch veld, Notitie aan TAW 3.
- [12] Tang, W.H., 1979, Probabilistic Evaluation of Penetration Resistances, Jrnl. Geot. Eng. Div. ASCE, GT10.
- [13] Calle, E.O.F., 1987, Toepassingen van Stochastiek en Statistiek in de Grondmehahnica, Syllabus Postdoctorale Cursus "Moderne Ontwikkelingen in de Grondmechanica", PATO Delft.

## 7. LIJST VAN GEBRUIKTE SYMBOLEN

De gebruikte symbolen in dit rapport worden in de tekst verklaard daar waar ze geïntroduceerd worden. De meest gebruikte symbolen worden in onderstaande lijst weergegeven.

$a_0$	autocorrelatieparameter voor correlatieafval op $\tau=0$ (nugget effect)
$\hat{a}_0$	meest aannemelijke (max. likelihood) schatter voor $a_0$
$\underline{a}$	vector met trendcoëfficiënten $a_1, a_2, \dots, a_6$
$\hat{\underline{a}}$	schatter voor $\underline{a}$ , berekend m.b.v. kleinste kwadraten methode
A	autocorrelatieparameter (komt overeen met $1/D^2$ )
$\underline{b}$	vector met (orthonormale) trendcoëfficiënten $b_1, b_2, \dots, b_6$
$\hat{\underline{b}}$	schatter voor $\underline{b}$
C	matrix voor normaalvergelijkingen in regressieprobleem
D	autocorrelatieparameter "fluctuatieschaal"
$\tilde{D}$	schatter voor D
$\hat{D}$	meest aannemelijke schatter voor D
$D(x,y)$	expressie voor gemiddelde trend
$\exp()$	exponentfunctie, gebaseerd op natuurlijk grondtal
$\bar{e}_i$	orthonormale polynomen, gebruikt bij trendanalyse
$E[ ]$	operator "verwachting"
G	bovendiaagonaalmatrix (decompositie van matrix C)
$G^T$	getransponeerde van matrix G (rij-kolomverwisseling)
$P(A)$	kans op een gebeurtenis A
$P(A B)$	conditionele kans op een gebeurtenis a, gegeven dat B optreedt
$\bar{q}_c$	over een bepaalde diepte gemiddelde conusweerstand
r	berekende correlatie tussen (twee) stochasten
$S_v$	standaardafwijking van stochast v, berekend uit steekproef
$\bar{v}$	rekenkundig gemiddelde waarde van realisaties van stochast v in een steekproef

bladnummer : - 49 -  
ons kenmerk: C0266483/14  
datum : 1987-07-28

- $V_v$  variatiecoefficient van stochast  $v$  ( $=\sigma_v/\mu_v = S_v/\bar{v}$ )
- $x, y, z$  ruimtelijke variabelen
- $\underline{x}$  vectornotatie voor ruimtelijke variabelen
- $\gamma(\tau)$  variogramwaarde als functie van "lag"  $\tau$
- $\delta(\underline{x})$  stochastische veldvariabele (incl. event. trend)
- $\delta_i$  waarneming aan stochastisch veld  $\delta(\underline{x})$  in punt  $\underline{x}_i$
- $\Delta x, \Delta y$  afstanden in  $x$  en  $y$  richting tussen meetpunten in het terrein
- $\phi()$  Gaussische (normale) kansverdelingsfunctie met gemiddelde 0 en standaardafwijking 1
- $\rho_\omega(\tau)$  autocorrelatiefunctie van stochastisch veld  $\omega()$
- $\rho_{ij}$  correlatie tussen stochastische veldvariabelen in de punten  $\underline{x}_i$  en  $\underline{x}_j$
- $\sigma_v, \sigma(v)$  standaardafwijking van stochast  $v$
- $\sigma_{res}$  residuele standaardafwijking van waarnemingen, na trend-eliminatie
- $\tau$  afstand ("lag") in autocorrelatiefunctie of variogram
- $\mu_v, \mu(v)$  verwachtingswaarde van stochast  $v$
- $\omega(\underline{x})$  veldvariabele van homogeen stochastisch veld

Appendix 1: Bewijs dat Tabba & Yong procedure (vergl. 2.17 en 2.18)  
 leidt tot zuivere schatters voor de fluctuatieschaal.

We beschouwen de verwachtingswaarde van  $R_i(\bar{D})$  voor een vaste waarde van  $\bar{D}=D'$ :

$$\begin{aligned} E[R_i(\bar{D}) |_{\bar{D}=D'}] &= \Sigma E[(\omega_i - r(D', \tau_{ij})\omega_j)^2] \\ &= \Sigma E[(\omega_i^2 - 2r(D', \tau_{ij})\omega_i\omega_j + r^2(D', \tau_{ij})\omega_j^2)] \\ &= \Sigma [1 - 2r(D', \tau_{ij})\rho_{ij} + r^2(D', \tau_{ij})] \end{aligned} \quad (aI.1)$$

Deze (conditionele) verwachtingswaarde wordt, onder variatie van  $D'$ , minimaal indien  $r(D', \tau_{ij}) = \rho_{ij}$ , voor elk van de voorkomende indices  $ij$ . Dit kan alleen als  $D'=D$ , de exacte waarde van de autocorrelatieparameter. Kennelijk geldt dus voor elke vaste waarde  $D'$  dat:

$$E[R_i(\bar{D} |_{\bar{D}=D'})] \geq E[R_i(\bar{D} |_{\bar{D}=D})] \quad (aI.2)$$

De waarde  $\bar{D}$ , zoals die bij de minimalisering van (2.17) wordt gevonden is afhankelijk van de stochasten  $\{\omega_i\}$  en daarom zelf ook een stochast.

De verwachtingswaarde van  $R_i(\bar{D})$ , ermee rekening houdend dat  $\bar{D}$  een stochast is kan worden geschreven als:

$$E[R_i(\bar{D})] = \int_{-\infty}^{\infty} E[R_i(\bar{D} |_{\bar{D}=D'})] p_{\bar{D}}^-(D') dD' \quad (aI.3)$$

waarin  $p_{\bar{D}}^-( )$  de (marginale) kansdichtheidsfunctie van  $\bar{D}$  is. Omdat deze positief is volgt met behulp van (aI.2):

$$E[R_i(\bar{D})] \geq E[R_i(\bar{D} |_{\bar{D}=D})] = E[R_i(D)] \quad (aI.4)$$

Omdat  $\bar{D}$  de eenduidige uitkomst is van een procedure die gebaseerd is op de voorwaarde (2.18), geldt dat:

$$R_i(\bar{D}) \leq R_i(D) \quad (aI.5)$$

en omdat linker en rechterlid in (aI.5) beide stochasten zijn die afhankelijk zijn van dezelfde verzameling  $\{\omega_i\}$  volgt dat:

bladnummer : - 51 -  
ons kenmerk: C0266483/14  
datum : 1987-07-28



GRONDMECHANICA  
DELFT

$$E[R_i(\bar{D})] \leq E[R_i(D)] \quad (\text{aI.6})$$

Uit (aI.4) en (aI.6) volgt dat noodzakelijk  $E[R_i(\bar{D})] = E[R_i(D)]$ . We concluderen dat als de stationariteitsconditie leidt tot een eenduidig bepaalde oplossing  $\bar{D}$ , verondersteld kan worden dat de verwachtingswaarde hiervan bij benadering gelijk moet zijn aan  $D$ .

De Technische Adviescommissie voor de Waterkeringen werd door de Minister van Verkeer en Waterstaat ingesteld.

De commissie adviseert de minister omtrent alle technisch-wetenschappelijke aspecten die van belang kunnen zijn voor een doelmatige constructie en het onderhoud van waterkeringen dan wel voor de veiligheid van door waterkeringen beschermde gebieden.

