

POSTFAȚĂ sau DECALOGUL GEOSTATISTIC

După ce am terminat de scris această carte am încercat să simulez o *amnezie geostatistică* și în această stare am *citit-o*. Concluzia a fost că GEOSTATISTICA ESTE UN MONSTRUOS HĂȚIȘ METODOLOGIC din care nu poți să ieși fără ajutor.

Povestea cu *pictura* îndulcește puțin efortul dar nu te ajută să-ți însușești *metodologia abstractă* căreia îi sunt supuse *datele* până sunt transformate într-o *secțiune litologică*, o *hartă cu izolinii* sau o *rețea de monitoring*. O uiți imediat ce dai de *normalitate*, *staționaritate*, *ergodicitate*, *variograme*, *kriging* și altele similare.

Ce ar trebui să nu uiți pentru a reuși să ieși cu bine din acest labirint metodologic?

Câte sunt poruncile ce trebuie respectate pentru a merge pe calea cea bună?

Decalogul geostatistic și soluțiile pe care vi le propun vă vor conduce cu siguranță la estimarea optimă a structurilor spațiale(E.S.S.).

1. Să prelucrezi numai selecții de date cu distribuție NORMALĂ!

Dacă selecția de date nu are o distribuție *normală* să o normalizezi înainte de prelucrare!

2. Să nu te lași atras de VALORILE EXTREME!

Dacă există *valori extreme* care afectează reprezentativitatea datelor, să le elimini!

3. Să prelucrezi selecții de date OMOGENE STATISTIC!

Dacă selecția de date este *neomogenă statistic*, să o separi în grupuri omogene!

4. Să aplici modelele geostatistice numai într-o ambianță STAȚIONARĂ!

Dacă identifici componenta *nestaționară*, să o separi de cea *staționară* și să o prelucrezi separat!

5. Să nu studiezi distribuția spațială a unei caracteristici fără să realizezi o REPRESENTARE GRAFICĂ sugestivă a datelor disponibile!

Dacă datele de care dispui îți sunt furnizate sub formă tabelară, *reprezentările grafice* pe care trebuie să le realizezi sunt: harta punctuală, harta indicatoare, diagrama de variabilitate și diagrama de continuitate!

6. Să stabilești LEGEA DE VARIAȚIE SPAȚIALĂ a caracteristicii studiate numai pe baza valorilor disponibile!

Dacă nu ai suficiente date pentru a identifica *legea de variație spațială*, să nu le forțezi să respecte o lege aleasă arbitrar!

7. Să testezi ANIZOTROPIA STRUCTURII SPAȚIALE studiate cu ajutorul legii de variație spațială!

Dacă *structura spațială este anizotropă*, să stabilești direcția și raportul de anizotropie!

8. Să validezi MODELUL LEGII DE VARIAȚIE SPAȚIALĂ (modelul de variogramă) stabilit!

Dacă erorile de estimare introduse de *modelul legii de variație spațială (modelul de variogramă)* stabilit depășesc limita admisă, trebuie să verifici dacă ai respectat primele 7 porunci sau să cauți alte metode pentru a identifica legea de variație spațială!

9. Să realizezi ESTIMAREA SPAȚIALĂ la gradul de detaliu adecvat datelor disponibile și obiectivului cercetării!

Dacă dispui de puține date trebuie să realizezi *estimări spațiale* de tip *global*, iar dacă dispui de multe date poți să realizezi *estimări punctuale sau zonale*.

10. Să calculezi EROAREA care afectează fiecare valoare estimată!

Dacă *eroarea* calculată depășește limita admisă să completezi datele disponibile cu date suplimentare obținute dintr-o rețea de cercetare proiectată adecvat!

Este obligatorie integralitatea respectării acestor 10 porunci pentru a obține estimarea optimă a structurilor spațiale.

Porunca a 8-a este cel mai greu de respectat și vă pune deseori în situația să reluați prelucrările de la început.

Nu ezitați să o luați de la început dacă rezultatul nu vă satisface!

Timpul necesar reluării unor prelucrări este din ce în ce mai scurt pe măsură ce numărul reluărilor este mai mare. Nu exagerați!

Câștigul de precizie obținut prin reluarea calculelor s-ar putea să fie semnificativ!

Nu uitați, modelele geostatistice sunt cele mai eficiente pentru *estimarea structurilor spațiale!*

După ce veți citi și partea a doua a *Geostatisticii aplicate* (S.S.S), veți vedea că modelele geostatistice sunt tot atât de performante și pentru *simularea structurilor spațiale!*

GEOSTATISTICĂ APLICATĂ

ESTIMAREA STRUCTURILOR SPAȚIALE

CUPRINS

Prefață	v
Ce este Geostatistica?	1
<i>Daniel Scrădeanu</i>	
Artă sau știință?.....	1
Terminologie.....	3
Cele trei vârste ale Geostatisticii.....	3
Obiectivul Geostatisticii aplicate	4
Etapele estimării structurilor spațiale.....	5
1. Analiza variabilității	7
<i>Daniel Scrădeanu</i>	
1.1. Analiza variabilității globale.....	7
<i>A1. Analiza unei repartiții normale și omogene</i>	8
<i>A2. Analiza unei repartiții lognormale și omogene</i>	13
<i>A3. Analiza unei repartiții lognormale și neomogene</i>	16
1.2. Normalizarea repartiției selecțiilor de date	20
<i>A4. Normalizare redusă</i>	22
<i>A5. Normalizare generalizată</i>	26
1.3. Analiza variabilității spațiale	28
1.3.1. Reprezentarea grafică.....	28
<i>A6. Harta punctuală</i>	31
<i>A7. Harta simbolică și indicatoare</i>	33
<i>A8. Diagrama de continuitate</i>	36
<i>A9. Diagrama de variabilitate</i>	39
1.3.2. Analiza parametrică a datelor	42
<i>A10. Covarianța, corelograma și variograma</i>	45
<i>A11. Intercovarianța, intercorelograma și intervareograma</i>	48
1.3.3. Analiza staționarității	54
<i>A12 Analiza staționarității cu ajutorul variogramei</i>	62
<i>A13. Eliminarea nestaționarității</i>	66
1.3.4. Analiza variografică.....	70
<i>A14. Variograma indicatoare</i>	87
<i>A15. Variograma omnidirecțională</i>	93
<i>A16. Analiza anizotropiei structurilor spațiale</i>	97
<i>A17. Validarea modelului de variogramă</i>	99
1.3.5. Analiza fractală	104
<i>A18. Calculul dimensiunii fractale a unei structuri de creștere</i>	109
<i>A19. Analiza distribuției fractale</i>	110

2. Estimarea distribuției spațiale	114
<i>Daniel Scrădeanu</i>	
2.1. Estimarea globală	116
<i>A20. Calculul valorii medii prin declustering celular</i>	<i>118</i>
2.2. Estimarea punctuală	121
2.2.1. Kriging punctual ordinar	124
<i>A21. Realizarea secțiunilor litologice.....</i>	<i>142</i>
<i>A22. Evaluarea distribuției unei variabile staționare</i>	<i>152</i>
2.2.2. Kriging punctual universal.....	155
<i>A23. Evaluarea distribuției unei variabile nestaționare.....</i>	<i>159</i>
2.3. Estimarea zonală	162
<i>A24. Calcul de rezerve</i>	<i>167</i>
<i>A25. Inițializarea parametrică a modelelor numerice</i>	<i>172</i>
3. Erorile estimării distribuției spațiale.....	176
<i>Daniel Scrădeanu și Roxana Popa</i>	
3.1. Calculul erorilor	176
<i>A26. Distribuția spațială a erorilor de estimare</i>	<i>190</i>
3.2. Reducerea erorilor.....	194
<i>A27. Optimizarea rețelei de monitoring</i>	<i>202</i>
4. Ghid de utilizare a pachetului de programe Geo-EAS	206
<i>Roxana Popa</i>	
4.1. Fișierele de date	207
4.1.1. Structura fișierelor.....	207
4.1.2. Convenții pentru extensia fișierelor	208
4.2. Ecrane interactive.....	208
4.3. Structura subprogramelor.....	210
4.4. Opțiuni comune.....	210
4.5. Grafică Geo-EAS.....	211
4.6. Etapele unui studiu geostatistic și programul Geo-EAS.....	211
4.7. Principalele subprograme Geo-EAS	212
4.7.1. Stat1	212
4.7.2. Vario	213
4.7.3. Xvalid	215
4.7.4. Krige	216
4.8. Subprograme utilitare Geo-EAS	218
4.8.1. Dataprep	218
4.8.2. Trans	220
4.8.3. Scatter	221
4.8.4. Xygraph.....	221
4.8.5. Postplot	222
4.8.6. Conrec.....	222
4.8.7. View.....	223
4.8.8. Hpplot	223
Postfață sau Decalogul Geostatistic	224
<i>Daniel Scrădeanu</i>	
Bibliografie.....	226

CE ESTE GEOSTATISTICA?

Paul Klee: “*Arta nu redă vizibilul, ci face vizibil.*”

Artă sau știință?

Problema fundamentală a cercetării geologice este să facă *vizibil* ceea ce este de cele mai multe ori ascuns observației directe.

Atât cercetarea geologică *fundamentală*, consacrată structurilor adânci ale Terrei, mecanismelor declanșării cutremurelor sau genezei zăcămintelor cât și cea *aplicativă*, dedicată exploataării zăcămintelor și protecției mediului geologic, trebuie să rezolve problema vizualizării rezultatului diverselor fenomene naturale, în scopul înțelegerii acestora.

Limbajul imaginilor vizuale și limbajul conceptelor nu sunt decât două ipostaze ale multiplului proces de cunoaștere și comunicare a omului cu natura. De-a lungul secolelor *arta* și *știința* au înaintat diferit în încercările lor de a înțelege *UNIVERSUL*.

Prin *artă* omul angajează un dialog permanent cu aparența naturii receptată printr-o variabilitate extraordinară a stimulilor senzoriali (vizuali, auditivi, tactili, olfactivi etc.). *Arta* fabrică din acești stimuli un alt fel de natură: o lume care prezintă aspectul *realității* și care se află totuși la dispoziția spiritului; o lume sensibilă, pătrunsă pe de-a-ntregul de intenții umane și construită în așa fel încât să realizeze perfectul acord al *conținutului* cu *forma*.

Artiștii s-au concentrat pe lucrurile cele mai frumoase și obișnuite: peisaje, portrete și viziuni proprii; *arta* își etala frumusețea.

Savanții gravitează în jurul polului opus, concentrându-se asupra experimentelor de laborator simplificate artificial, unde observatorul este redus la o imposibilă variabilă controlată. Eleganța de elită a științei se ascunde în spatele ecuațiilor. Chiar și geometria, cea mai artistică dintre ramurile matematicii, a devenit mai mult un exercițiu logic decât vizual.

Geostatistica, în demersul de clarificare a proceselor geologice, apelează în mod egal la *limbajul imaginilor* și al *conceptelor* la fel ca *Pictura* care degajează adevărul de formele vane și amăgitoare ale acestei lumi imperfecte pentru a-l îmbrăca într-o formă mai pură și mai elevată creată de spiritul însuși.

Departate de a fi simple aparențe, formele artei cuprind mai multă realitate și mai mult adevăr decât existențele fenomenale ale lumii reale. “Spiritul răzbate mai greu prin învelișul dur al naturii și al vieții obișnuite decât prin operele de artă” (Hegel).

Similaritățile între *Geostatistică* și *Pictură* fac interesantă abordarea arsenalului metodologic al Geostatisticii din punctul de vedere al pictorului așezat în fața unui peisaj pe care vrea să-l transpună într-un tablou.

Ar reuși un *pictor*, utilizând programul Geo-EAS în locul pensulelor și vopselelor, să immortalizeze epava de la Vama Veche îmbrățișată de zbuciumul valurilor ?

Ar reuși un *geolog* să realizeze o secțiune litologică pe baza datelor din forajele de explorare utilizând pensulele și vopselele pictorului ?

Dacă pentru pictor ar exista probleme mari ridicate de utilizarea programului și a modelelor geostatistice, geologul ar reuși cu pensule și vopsele să-și realizeze secțiunea fără mari dificultăți. Tehnica picturii se potrivește ca o mânășă pentru

“plăsmuirea” secțiunilor geologice mai ales atunci când datele sunt puține și nu se prea “leagă”.

Pictura operează cu *limbajul imaginilor* și al *conceptelor* în procesul de cunoaștere a naturii și are ca elemente constitutive, pe de o parte *desenul* și *forma*, iar pe de alta - *lumina* și *culoarea*. Aceste elemente cu aceleași materializări dar cu funcții și finalități diferite le regăsim atât în tabloul pictorului cât și în secțiunea geologică elaborată pe baza modelelor geostatistice:

- *forma* lucrurilor interesează foarte mult conduita noastră practică, ea ne relevă aspectul tangibil al lucrurilor și se realizează prin *desen*;
- *culoarea* se adresează intens sentimentelor noastre, foarte diversificate în cazul picturii, reduse în cazul geologiei la sentimentul de *certitudine* sau *incertitudine* asociat formei realizate.

Începutul *picturii* este un gest imitativ: oamenii sălbatici fac în aer, pe nisip sau pe peretele peșterii crochiul rapid al animalului într-o anumită mișcare. În această imitație există două lucruri: un *limbaj abstract* al științei care are un rol *indicativ* și *limbajul viu* și concret al artei cu rol *descriptiv*.

La început, *desenul* nu este decât un mijloc de a indica, de a figura un lucru, altfel spus, un *limbaj grafic*. Ulterior, își fac apariția tendința descriptivă, efortul de a domina schema intelectuală și a ajunge la realitatea concretă a obiectului, la reprezentarea sintetică a ansamblului, la punerea în valoare a detaliilor, la construcția precisă. Realismul vizual luptă împotriva schematismului logic.

Percepția vizuală a naturii este o cucerire a vârstei mature, a civilizației. Ea presupune o analiză și o filtrare a stimulilor senzoriali.

Grecii erau foarte legați de contururile frumoase; ei vedeau pictura cu ochiul sculptorului de statui și, chiar când pictau, căutau mai presus de orice *liniile* care marchează limitele corpului în spațiu, acordau *desenului* rolul preponderent. Ingres spunea că *desenul* cuprinde în el trei sferturi și jumătate din ceea ce constituie *pictura*.

Desenul este descrierea formelor și el deține supremația în realizarea hărților și secțiunilor geologice. A desena nu înseamnă doar a reproduce contururi: *linia care reprezintă conturul formelor este o abstracție a modelului*. Marea calitate a desenului constă în felul în care reușește să redea expresia vieții proprii a formei prin acest contur. *Desenul nu este doar o linie care delimitează un corp ci este expresia tuturor formelor lui, planul modelat al aceluia corp*.

Forma unui corp nu este niciodată prea precisă. Un mare maestru, prin simplificare, trebuie să scoată efectele cele mai puternice pentru desenul realizat. Detaliile “sunt niște mici pisălogi pe care trebuie să-i punem la respect”. *Desenul* pentru sine însuși este analiza plastică a lucrurilor, cu *subordonarea culorii față de formă*. Desenatorul pur este un filosof care extrage chintesența și “hăituiește întotdeauna culoarea”.

Geostatistica operează asupra *formelor* de vizualizare a obiectului de studiu iar *culorii* îi acordă rolul de *mesager al erorilor* cu care sunt construite aceste forme.

Geostatistica aplicată este consacrată prezentării, sub aspect operațional, a unei *practici* bazate pe un ansamblu de *modele* și *tehnici* elaborate de-a lungul anilor, în principal la Centrul de Geostatistică și Morfologie Matematică al Școlii de Mine de la Fontainebleau, pentru *descrierea*, *studierea* și *estimarea* celor mai variate fenomene cu dezvoltare spațială: ”il peut s’agir de cellules cancéreuses ou de forêts, de la structure d’une roche ou celle d’un alliage métallique, d’un gisement de pétrole ou d’un revêtement de route, de pollution ou de météorologie, de cartographie sous-

marine ou de prospection géophysique etc”** (G. Matheron, "Estimer et choisir", 1978).

Această metodologie își are originea într-o problemă legată de cercetarea geologică: *evaluarea rezervei zăcămintelor*. Evaluarea rezervelor zăcămintelor a servit ca stand de încercări și validare pentru majoritatea metodelor *Geostatisticii*.

Modelele geostatistice oferă una din căile cele mai performante pentru descrierea structurii spațiale a fenomenelor naturale, caracterizate printr-o extraordinară variabilitate *spațială* și *temporară*, variabilitate determinată de complexitatea fizico-chimică a mediului și instabilitatea proceselor climatice.

Terminologie

Neologismul "Géostatistique" a apărut în 1962 și prin prefixul "geo" atrăgea atenția asupra luării în considerare a *repartiției spațiale* a valorilor prelucrate. Până la acel moment, titlurile articolelor care prezentau metodele ce se constituiau în arsenalul *geostatisticii primei vârste* făceau referiri directe la metode *statistice* care în forma lor clasică erau insuficiente pentru studiul zăcămintelor diseminate.

După 1970, când metodele geostatistice se dezvoltă în direcția lor proprie și se diferențiază din ce în ce mai mult de *statistica clasică*, termenul de *geostatistică* generează confuzii. În mod curent, în România, *Geostatistica* este percepută de cei neavizați într-o optică strict statistică. Traducerea termenului în alte limbi (*Geostatistics* în limba engleză, *Geostatistică* în limba română) a amplificat și mai mult confuziile. G. Matheron introduce în 1978 ("Estimer et choisir") pentru a înlocui termenul *geostatistică* expresia *modele topo-probabiliste* care are dublul avantaj de a distanța metoda de domeniul geologic, prin eliminarea prefixului "geo" și de a sublinia caracterul mai mult probabilist decât statistic al metodei.

Cele trei vârste ale Geostatisticii

Analizată atât din punct de vedere al evoluției teoretice cât și al domeniului de aplicare, *Geostatistica* a parcurs trei perioade distincte până în prezent.

Prima vârstă a Geostatisticii este de inspirație strict minieră. Lucrările miniere din Africa de Sud, care au determinat primele cercetări în acest domeniu, aparțin cercetătorilor *H.S.Sichel*, *D.G.Krige*, *H.J. de Wijs*. Completarea metodologiei statistice clasice a fost determinată de noile exigențe impuse de cercetarea zăcămintelor foarte diseminate (aur, uraniu, nichel, cupru, etc.).

La nivel teoretic, formalismele matematice specifice acestei perioade se plasează în cadrul unei *legi de distribuție dată*, modelul lognormal având o epocă de aur în anii 1950-1970 (*Sichel H.S.*, 1966).

La nivel practic, mijloacele de calcul sunt rudimentare, motiv pentru care lucrările de specialitate abundă în *formule de aproximare*, *curbe* și *abace* utilizate pentru a evita reluarea unor calcule laborioase.

A doua vârstă a Geostatisticii corespunde perioadei 1965-1970, perioadă în care modelele statistice sunt abandonate. Se elaborează modele care nu implică legea de repartiție (*Geostatistica lineară*). În același timp, se încearcă extinderea ipotezelor

* "poate să fie celule cancerigene sau păduri, structura unei roci sau a unui aliaj, un zăcământ de petrol sau îmbrăcăminte de șosele, poluare sau meteorologie, cartografie submarină sau prospecțiuni geofizice."

de lucru prin dezvoltarea *Geostatisticii nestaționare*, apoi a *Geostatisticii neliniare*. În această perioadă se dezvoltă: *simularea condițională și necondițională, ansamblele aleatoare*. Toate aceste procedee noi sunt imediat aplicate datorită remarcabilei ameliorări a mijloacelor de calcul. Progresele teoretice și metodologice ale geostatisticii vârstei a doua se concretizează abia în 1978 printr-o singură lucrare de sinteză - *Estimer et Choisir* - elaborată de G.Matheron. În paralel cu aceasta, apar o mulțime de cărți de geostatistică cu pronunțat caracter *informatic* care pun la dispoziția practicienilor instrumentele aplicării metodelor clasice sau mai puțin clasice ale Geostatisticii.

Geostatistica *vârstei a treia* este în plină desfășurare și, în contextul informatic confortabil al sfârșitului de mileniu, se dezvoltă în direcții din ce în ce mai diverse. Atât la nivelul domeniului de aplicare cât și al domeniului teoretic se remarcă reconsiderarea legii de repartiție, nu ca un regres teoretic ci ca o completare necesară în special *simulării stohastice*.

Obiectivul Geostatisticii aplicate

Adaptată studiului distribuției spațiale pentru orice *variabilă regionalizată*, *Geostatistica* și-a conturat până în prezent câteva utilități clasice în domeniul cercetării geologice.

Geostatistica aplicată are ca obiectiv clarificarea metodologiei specifică modelelor geostatistice care permit rezolvarea problemelor clasice ale cercetării geologice:

a) *probleme curente* ale cercetării geologice:

- proiectarea rețelelor de explorare a zăcămintelor de orice tip;
- calculul rezervelor pentru zăcămintele de substanțe minerale utile;
- cartografierea automată a parametrilor geologici cantitativi (grosimi, adâncimi ale reperelor stratigrafice și litologice) și calitativi (litologie, vârstă etc.);
- simularea litofacială și parametrică a rezervoarelor (acvifere, zăcăminte de petrol);

b) *probleme noi* impuse de accentuarea laturii ecologice a cercetării geologice :

- proiectarea rețelelor de monitorizare integrată a parametrilor de mediu;
- cartografierea distribuției spațiale a parametrilor ambientali;

c) *probleme operaționale* conexe modelării matematice a proceselor geologice:

- inițializarea parametrică a modelelor matematice;
- calarea modelelor matematice de simulare a proceselor geologice;
- evaluarea incertitudinii rezultatelor simulării proceselor geologice prin modelare matematică.

Utilizarea modelelor geostatistice în *proiectarea rețelelor de explorare a zăcămintelor* permite ameliorarea pe parcursul execuției lucrărilor a geometriei rețelelor de explorare. Practica arhaică de îndesire a rețelelor de explorare după standarde care nu țin seama de particularitățile structurilor spațiale ale caracteristicilor studiate este abandonată de mult timp. Modelele geostatistice oferă instrumentul identificării geometriei optime a rețelelor de cercetare adaptată configurației spațiale a variabilelor studiate.

Evaluarea corectă a rezervelor zăcămintelor, în condițiile variabilității pronunțate a conținuturilor, presupune cunoașterea valorii erorilor de estimare. Modelele topo-probabiliste oferă o metodologie de cuantificare a erorilor de estimare și implicit a riscurilor economice asumate la exploatarea rezervelor evaluate.

Cartografierea automată a caracteristicilor de orice tip (*cantitative*: cota acoperișului sau grosimea unui strat, conținutul în aur, porozitatea unui acvifer etc. sau *calitative*: tip litologic, grad de alterare etc.) dispune în prezent de numeroase programe ce aplică diferite *metode de interpolare*. Modelele geostatistice oferă cea mai performantă și flexibilă metodă de interpolare ("*kriging*"), sensibilă la particularitățile structurii spațiale ale variabilei cartografiate.

Prin intermediul *simulărilor condiționate* realizate cu modelele *topo-probabiliste (geostatistice)*, datele litologice obținute din cartările aflorimentelor și din forajele de explorare sunt utilizate pentru generarea repertoriului de litofaciesuri corespunzător unui anumit grad de cunoaștere a domeniului spațial cercetat.

Modelele geostatistice sunt singurele care permit analiza *corelațiilor spațio-temporale multivariate* implicate în *proiectarea rețelelor de monitorizare integrată a parametrilor ambientali*.

Distribuția spațială corectă a parametrilor fizico-chimici pentru modelele de simulare numerică a proceselor geologice (dinamice sau staționare) asigură *reprezentativitatea* acestora pentru procesele geologice cercetate. Modelele geostatistice sunt cele care realizează *inițializarea parametrică* a modelelor matematice pentru: zăcăminte de substanțe minerale solide, zăcăminte de petrol, acvifere.

Testarea *reprezentativității* modelelor matematice prin *calarea și cuantificarea incertitudinii rezultatelor simulărilor* obținute se realizează în mod riguros utilizând *modelele geostatistice*. *Modelele geostatistice* permit exprimarea într-un mod sugestiv a diferențelor dintre parametrii măsurați și cei evaluați oferind soluția creșterii *reprezentativității* modelelor matematice.

Etapele estimării structurilor spațiale

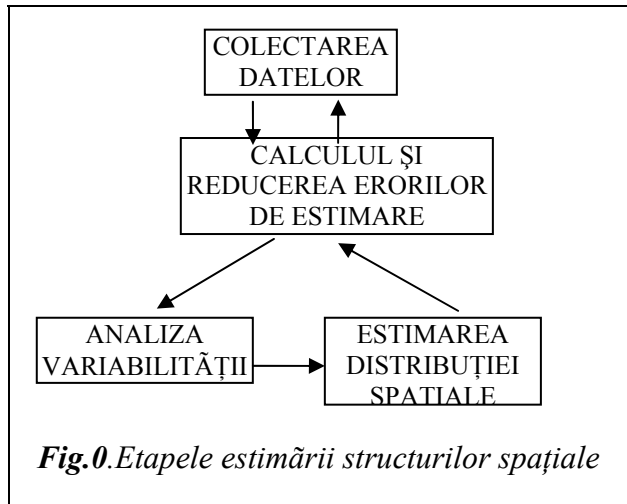


Fig.0. Etapele estimării structurilor spațiale

Estimarea structurilor spațiale (E.S.S.), materializată de regulă sub forma *hărților și secțiunilor geologice*, este rezultatul unor prelucrări laborioase a căror succesiune trebuie respectată cu strictețe. În condițiile unor volume de date semnificative, realizarea *E.S.S.* este de neconceput fără un calculator și un program performant. Programele care realizează astfel de estimări sunt structurate în module, fiecare dintre acestea realizând o

categorie de prelucrări specifice fiecărei etape. Numai cunoașterea conținutului fiecărei etape de prelucrare și a succesiunii în care acestea trebuie parcurse asigură succesul estimării structurilor spațiale cu ajutorul programelor.

Principalele etape ale *E.S.S.* (**Fig. 0**) sunt: *colectarea datelor, analiza variabilității caracteristicilor geologice, estimarea distribuției spațiale, calculul și reducerea erorilor de estimare.*

Colectarea datelor se realizează într-o rețea de explorare preliminară ale cărei caracteristici (geometrie, frecvență de prelevare etc.) se adaptează variabilității caracteristicii cercetate. Tehnologia de prelevare a probelor, de determinare a valorilor prelucrate, de stocare a acestora în baze de date și alte probleme ale colectării datelor sunt prezentate în detaliu în primul volum de INFORMATICĂ GEOLOGICĂ (1995). *Geostatistica aplicată* prezintă în detaliu ultimele trei etape.

Analiza variabilității caracteristicilor geologice este cea mai laborioasă etapă de prelucrare și are ca obiective identificarea *legilor globale de repartiție* a acestora, *normalizarea* lor și *găsirea legilor de variație spațială*. Alegerea legii de variație spațială a caracteristicii studiate nu este o operațiune simplă și pe lângă prelucrări cantitative presupune un simț special care se formează pe parcursul prelucrării datelor. De cele mai multe ori soluția nu este unică, testele utilizate conducând la “domenii de valabilitate”, astfel încât, cu aceleași date, zece geostatisticieni cu experiență pot alege zece legi de variație spațială egal valabile.

Estimarea distribuției spațiale poartă amprenta legii de variație spațială. Pe baza acestei legi, în această etapă, pentru o caracteristică oarecare se poate preciza în orice punct din interiorul domeniului studiat, care este *probabilitatea* ca în acel punct să fie prezentă variabila studiată și care este *cea mai probabilă valoare* a acesteia în acel punct. Există variante de estimare a distribuției spațiale în funcție de dimensiunea suportului pe care ea se realizează:

- *estimare globală*, dacă suportul estimării este întreg domeniul cercetat;
- *estimare punctuală*, dacă estimarea se realizează pentru un suport punctual;
- *estimare zonală*, dacă se estimează o valoare medie pentru o suprafață finită delimitată de un contur închis.

Calculul erorilor este ultima etapă de prelucrare în care se calculează eroarea care afectează fiecare variantă de estimare. *Eroarea minimă* de estimare a distribuției spațiale pentru un anumit grad de cunoaștere se obține prin utilizarea *kriging*-ului ca metodă de interpolare. Procedul de calcul al erorilor de estimare prin *kriging* permite optimizarea rețelelor de explorare. Etapa de calcul a erorilor mai este folosită și ca o primă etapă de filtrare a datelor colectate când sunt vizate doar erorile introduse de procedeele de prelevare a probelor și realizare a determinărilor.

Parcurgera ciclică a etapelor de prelucrare, în sensul indicat de săgeți (**Fig.0**), are ca obiectiv creșterea gradului de precizie a estimării structurilor spațiale până la eroarea impusă de obiectivul cercetării. *Reducerea erorilor* de estimare presupune completarea datelor primare prin prelevări de probe din domeniile spațiale în care densitatea punctelor de observație este redusă sau variabilitatea caracteristicilor ridicată.

1. ANALIZA VARIABILITĂȚII CARACTERISTICILOR GEOLOGICE

Analiza *variabilității* caracteristicilor geologice are ca obiectiv filtrarea tuturor *datelor* colectate din rețeaua de explorare preliminară în scopul pregătirii instrumentelor adecvate obținerii *desenului*.

Datele sunt constituite din coordonatele spațiale ale punctelor de observație, conținutul în aur determinat în proba analizată, vârsta sau litologia probei, conținutul în hidrocarburi etc.

Desenul poate reprezenta conturul zonei cu conținutul în aur mai mare decât conținutul minim exploatabil, limita de extindere a unei formațiuni de o anumită vârstă sau litologie, conturul unei zone poluate cu hidrocarburi etc.

Pentru *pictor* această etapă constă în contemplarea a peisajului și eventual realizarea unei *schițe*, în timp ce pentru *geostatistician* reprezintă începutul unei lungi serii de prelucrări care pleacă de la date și se finalizează sub forma unor *funcții de distanță*, inexpresive și neprietenoase, care trebuie să conțină esența desenului pe care vrea să-l realizeze.

Extraordinara *variabilitate* a caracteristicilor geologice analizate se manifestă sub două aspecte: *global*, exprimat prin distribuția valorilor caracteristicii în jurul unei valori centrale (*media* sau *mediana*) și *spațial*, exprimat prin variația *valorii* caracteristicii (v_i) în funcție de *poziția* în spațiul geometric de desfășurare a procesului în care este implicată (poziție exprimată de obicei prin coordonatele x_i, y_i, z_i , i fiind indicele punctului de observație în care este măsurată valoarea v).

1.1. ANALIZA VARIABILITĂȚII GLOBALE

Analiza *variabilității globale* a caracteristicilor geologice vizează asigurarea *reprezentativității* evaluărilor și se realizează prin: analiza *modului de distribuție*, a *eterogenității* și a *valorilor extreme* ale selecțiilor de date.

Metodologia evaluării *geostatistice a unei caracteristici este elaborată pentru distribuția normală a acesteia* și din acest motiv *neconcordanța dintre distribuția valorilor prelucrate și cea normală* conduce la *supraestimări* sau *subestimări* proporționale cu *gradul de asimetrie* al acestei distribuții. Prin urmare, *modul de distribuție* al valorilor unei caracteristici geologice în jurul *mediei* sau *mediane* de selecție influențează în mod determinant rezultatele prelucrărilor geostatistice.

Analiza modului de distribuție al frecvenței valorilor, în cazul pregătirii acestora pentru prelucrări geostatistice, poate conduce la două rezultate:

- ***repartiția valorilor este normală*** și în consecință prelucrarea lor prin modele geostatistice conduce la rezultate corect interpretabile;
- ***repartiția valorilor nu este normală***, în acest caz fiind necesară transformarea lor (*normalizarea*) în scopul eliminării erorilor introduse prin subestimări sau supraestimări.

Analiza eterogenității selecției de date disponibile, realizată de obicei prin *analiza dispersională multifactorială* (D.Scrădeanu, 1995), are ca obiectiv separarea selecției de date în funcție de dispersia diferențiată a valorilor, diferențiere determinată de factori fizico-chimici cu *acțiune divergentă*. Rezultatul analizei eterogenității selecției de date poate conduce la două variante de continuare a estimărilor geostatistice :

- **selecția de date este omogenă**, variantă în care estimările geostatistice se realizează asupra întregului set de date utilizându-se un singur model de variabilitate spațială ;
- **selecția de date este eterogenă**, variantă în care selecția de date trebuie separată în grupuri omogene, pentru fiecare din acestea identificându-se modele distincte de variabilitate spațială.

Analiza valorilor extreme ale selecțiilor de date asigură estimarea corectă a intervalului de încredere al parametrilor statistici prin corectarea valorilor exagerate ale dispersiei. Includerea *valorilor extreme* în prelucrare modifică semnificativ *dispersia de selecție*, conducând la creșterea artificială a gradului de incertitudine al evaluărilor statistice și geostatistice. Analiza valorilor extreme poate conduce și ea la două situații distincte:

- **valorile extreme se elimină** deoarece sunt puțin numeroase și din punct de vedere statistic nu sunt reprezentative pentru caracteristica studiată (sunt fie rezultatul unor erori de măsurare fie al unor variații bruște ce nu sunt definitorii pentru variabilitatea spațială a caracteristicii studiate);
- **valorile extreme nu se elimină** deoarece sunt suficient de numeroase pentru a putea forma o selecție de date căreia i se aplică metode specifice de prelucrare (P.Bomboe, 1979).

A1. Analiza unei repartiții normale și omogene

Ce mod de repartiție au grosimile depozitelor daciene măsurate în 124 de foraje (Tabelul 1.1) distribuite relativ uniform în Bazinul Dacic (Fig.1.1) ?

Rezolvare:

Compararea bazinelor de sedimentare, identificarea direcțiilor de transport a sedimentelor, zonarea acestor bazine, utilizează frecvent ca date primare *grosimile formațiunilor* depuse în anumite intervale de timp (Scrădeanu et.al.,1998). Bazinul Dacic, unitatea sedimentară lacustră din estul Paratethisului, a fost intens cercetat prin foraje, în legătură cu zăcămintele de petrol.

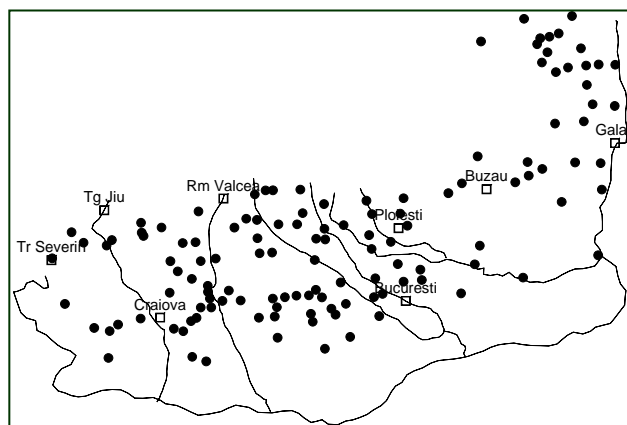


Fig.1.1. Distribuția celor 124 de foraje care au traversat formațiunile daciene

Stabilirea modului de repartiție al grosimii formațiunilor daciene se realizează în două etape :

- analiza grafică a modelului de distribuție ;
- testarea analitică a concordanței cu modelul de repartiție identificat.

Tabelul 1.1. Grosimi ale depozitelor daciene din Bazinul Dacic

Nr. crt.	Grosime g[m]	Nr. crt.	Grosime g[m]	Nr. crt.	Grosime g[m]	Nr. crt.	Grosime g[m]
1	602.47	32	555.46	63	551.54	94	244.11
2	607.20	33	406.81	64	539.97	95	83.55
3	927.18	34	750.22	65	622.09	96	192.70
4	450.37	35	649.77	66	638.18	97	216.03
5	662.22	36	528.71	67	383.54	98	235.44
6	737.94	37	627.29	68	559.45	99	367.14
7	799.48	38	496.16	69	588.81	100	503.28
8	617.00	39	536.19	70	575.85	101	410.55
9	391.45	40	780.16	71	612.05	102	310.32
10	499.71	41	453.90	72	571.67	103	543.78
11	532.44	42	492.62	73	683.04	104	485.56
12	675.76	43	325.98	74	824.11	105	478.52
13	859.18	44	726.91	75	643.87	106	330.93
14	584.43	45	698.84	76	428.93	107	259.91
15	371.32	46	707.51	77	354.17	108	15.27
16	395.34	47	632.66	78	521.33	109	547.64
17	482.04	48	517.69	79	580.10	110	315.68
18	928.00	49	506.85	80	525.01	111	425.30
19	716.81	50	432.54	81	690.71	112	510.45
20	563.46	51	471.50	82	439.71	113	399.20
21	567.54	52	475.00	83	178.82	114	274.13
22	597.83	53	446.83	84	414.27	115	320.89
23	514.06	54	421.65	85	443.27	116	299.09
24	358.55	55	304.80	86	464.47	117	293.21
25	340.51	56	162.71	87	436.13	118	387.51
26	457.43	57	403.02	88	349.70	119	345.15
27	118.71	58	467.99	89	362.88	120	252.24
28	379.52	59	460.95	90	267.21	121	226.13
29	764.08	60	417.98	91	204.98	122	143.42
30	668.83	61	489.08	92	280.74	123	335.78
31	593.28	62	655.88	93	287.09	124	375.44

Analiza grafică a modelului de distribuție se bazează pe *histograma (Fig.1.2.a)* și *diagrama de probabilitate (Fig.1.2.b)* ale grosimilor depozitelor daciene.

Analiza *histogramei* conduce la următoarele observații :

- *histograma are un singur modul* - indicație clară a caracterului *omogen* (din punct de vedere statistic) al selecției; consecința acestei observații este

că din punct de vedere statistic toți factorii care au determinat distribuția grosimilor depozitelor dacianului au acționat convergent ;

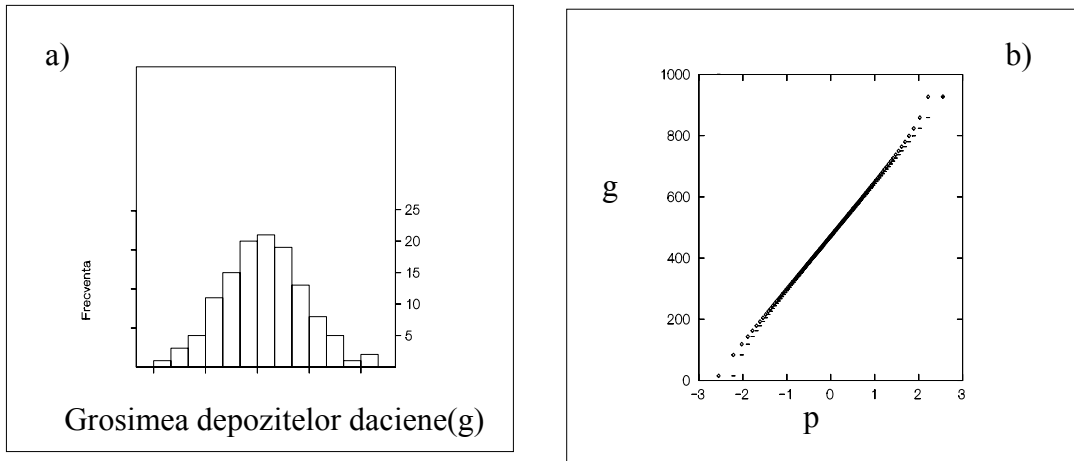


Fig.1.2. Histograma (a) și diagrama de probabilitate (b) ale grosimii depozitelor daciene (Bazinul Dacic)

- caracterul *simetric al histogramei* (coeficientul de asimetrie: $\beta_1 = 0,01$) sugerează o *distribuție normală* a valorilor; corectitudinea acestei ipoteze urmează să fie testată prin metode analitice .

Analiza *diagramei de probabilitate*, întocmită pentru repartiția normală, permite sesizarea următoarelor caracteristici :

- punctele ce reprezintă valorile grosimilor se *coliniarizează* - indicație clară a *distribuției normale* a acestora ;
- la extremitățile graficului se plasează un număr de 4 *valori extreme* (două la limita inferioară și două la cea superioară) a căror apartenență la selecție trebuie testată (printr-un test analitic; ex.: Chauvenet, Irwin, Romanovski etc.) .

Testarea analitică a concordanței repartiției valorilor cu modelul repartiției normale se realizează de cele mai multe ori cu testul χ^2 .

Aplicarea testului χ^2 se bazează pe histograma celor $n = 124$ valori de grosimi, histogramă care s-a realizat prin gruparea datelor pe $k = 13$ intervale, fiecare având amplitudinea valorică $\Delta = 75 m$.

Media aritmetică ponderată a celor 124 de valori este $m = 475,16 m$, abaterea lor standard $s = 176,24 m$ iar valorile centrale ale intervalelor de grupare (x_{ci}) și frecvențele absolute experimentale (n_i) sunt grupate în **Tabelul 1.2**.

Datele necesare testului χ^2 sunt parametrii statistici elementari și valorile din primele patru coloane din **Tabelul 1.2**.

Calculul frecvențelor teoretice (np_i), corespunzătoare repartiției normale (coloana a cincea din **Tabelul 1.2**), și al statisticii χ^2_{exp} se efectuează pe baza relațiilor (D.Scrădeanu,1995):

$$np_i = \frac{\Delta \cdot n}{s \cdot \sqrt{2 \cdot \pi}} \cdot \text{Exp} \left[-0,5 \cdot \left(\frac{x_{ci} - m}{s} \right)^2 \right] \quad \text{și} \quad \chi^2_{exp} = \sum_{i=1}^{i=k} \frac{(n_i - np_i)^2}{np_i}$$

Prin eliminarea intervalelor de grupare extreme, cu frecvențe absolute mai mici de doi (P.Bomboe,1979), se obține:

$$\chi^2_{\text{exp}} = 0.46$$

Pentru un risc al erorii de genul I, $\alpha = 1\%$, la un număr al gradelor de libertate $\nu = 7$ (după eliminarea celor trei intervale de grupare extreme) rezultă (din tabelul funcției χ^2):

$$\chi^2(0,99;7) = 1,23$$

Conform criteriului testului de concordanță χ^2 , deoarece :

$$(\chi^2_{\text{exp}} = 0,46) < (\chi^2(0,99;7) = 1,23)$$

rezultă că, din punct de vedere statistic, cu o eroare admisibilă $\alpha = 1\%$, distribuția grosimilor dacianului este conformă cu modelul repartiției normale.

Tabelul 1.2. Calculele asociate testului χ^2

Nr. interval	Lim. inf. interval	Lim. sup. interval	x_{ci}	n_i	np_i	$\frac{(n_i - np_i)^2}{np_i}$
1	0	75	15.3	1	0.70	0.13
2	75	150	115.2	3	2.62	0.06
3	150	225	191.0	5	5.74	0.10
4	225	300	265.4	11	10.37	0.04
5	300	375	339.6	15	15.66	0.03
6	375	450	411.9	20	19.74	0.00
7	450	525	485.7	21	21.02	0.00
8	525	600	560.2	19	18.74	0.00
9	600	675	633.8	13	14.04	0.08
10	675	750	704.7	8	9.01	0.11
11	750	825	783.6	5	4.55	0.04
12	825	900	859.2	1	1.96	0.47
13	900	975	927.6	2	0.78	1.91
$\chi^2_{\text{exp}} =$						0.46

COMENTARIU :

Grosimile depozitelor daciene, măsurate în cele 124 de foraje din Bazinul Dacic (Fig.1.1) au o repartiție normală.

Are importanță acest rezultat ?

Desigur, stabilirea acestuia (grosimile depozitelor daciene măsurate în cele 124 foraje au repartiție normală) ne asigură că estimările care se pot face pe baza grosimilor depozitelor daciene, prin metode geostatistice (grosimi medii, zonarea grosimilor, direcții de transport etc.), nu vor fi afectate de supraestimări sau subestimări.

Este justificat efortul de calcul necesar formulării acestei concluzii ?

Orice efort de calcul este mult mai puțin costisitor decât obținerea datelor care se prelucrează. Comod așezat în fața unei mașini de calcul și cu un program adecvat, care nu costă prea mult, în maximum 10 minute se realizează testele minime necesare precizării răspunsului la întrebarea aplicației.

Ce limitări sunt asociate acestei concluzii ?

Limitările sunt determinate de reprezentativitatea datelor disponibile și trebuie să medităm la cel puțin două întrebări în legătură cu acestea :

- *Cele 124 de grosimi disponibile, determinate pe locația celor 124 de foraje, sunt suficiente pentru a descrie variabilitatea globală a grosimii depozitelor daciene din arealul cercetat ?*
- *Grosimile formațiunilor daciene au o distribuție normală în Bazinul Dacic sau numai cele 124 de valori prelucrate îndeplinesc această condiție ?*

Metodologia statistică, corect aplicată, asigură reprezentativitatea rezultatelor obținute și nu pe cea a datelor utilizate.

Atenție deci la calitatea datelor prelucrate !!!

A2. Analiza unei repartiții lognormale și omogene

Ce distribuție a frecvențelor au grosimile depozitelor romaniene măsurate în 101 foraje (**Fig.1.3** și **tabelul 1.3**) distribuite relativ uniform în Bazinul Dacic ?

Rezolvare :

Analiza grafică a distribuției grosimilor depozitelor romanianului indică o asimetrie evidentă. *Histograma (Fig.1.4a)* indică o puternică asimetrie de stânga ($\beta_1=2,3$) iar în *diagrama de probabilitate normală (Fig.1.4b)* aranjamentul punctelor este de-a lungul unei curbe.

Morfologia histogramei și a diagramei de probabilitate a datelor originale sugerează o distribuție lognormală a acestora.

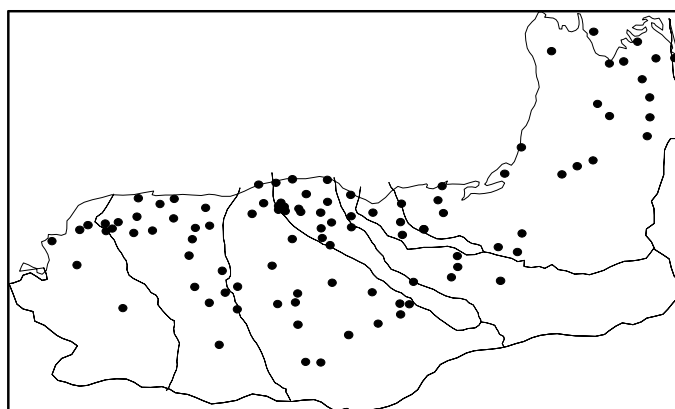


Fig.1.3. Distribuția celor 101 foraje.

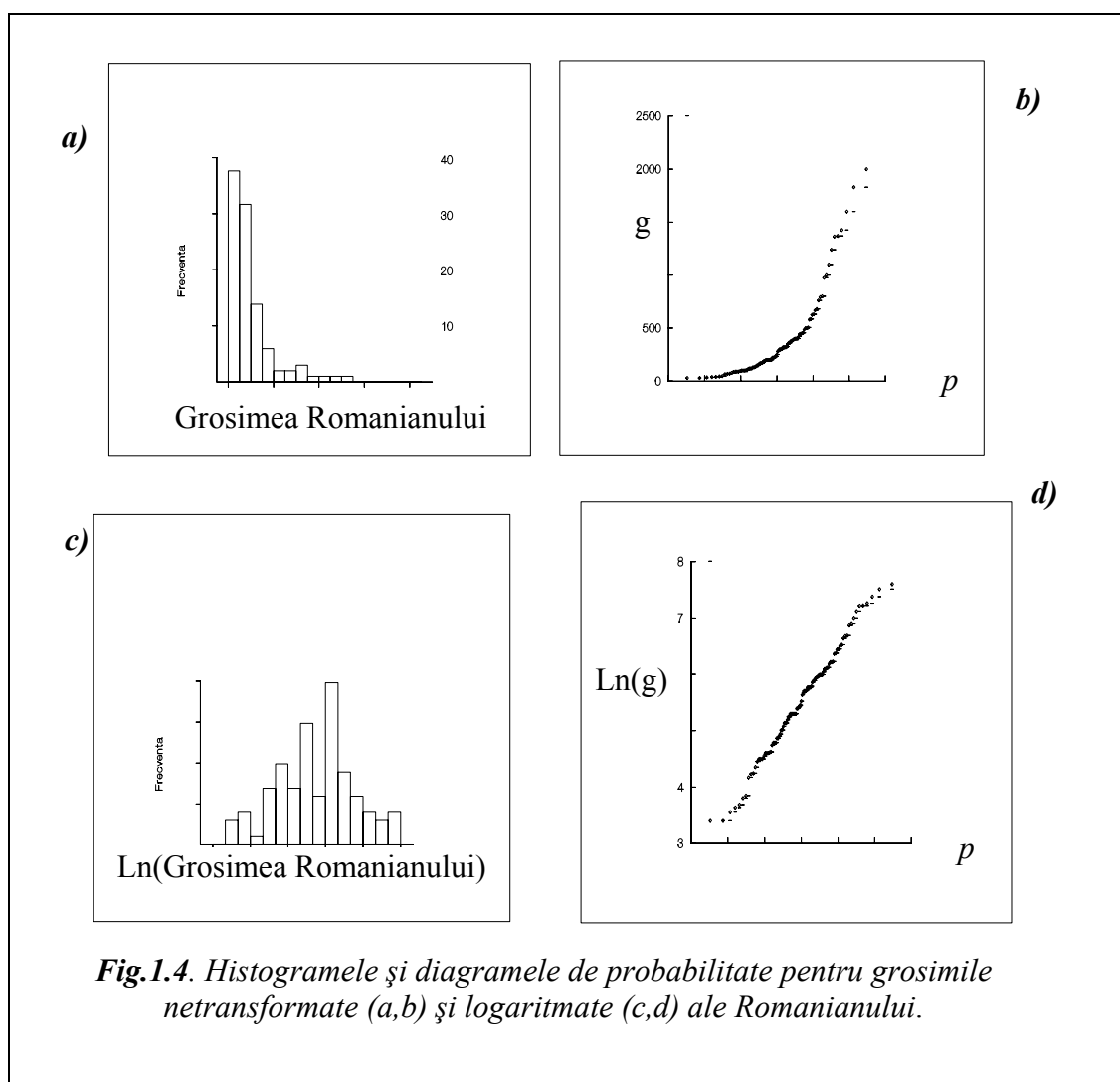


Fig.1.4. Histogramele și diagramele de probabilitate pentru grosimile netransformate (a,b) și logaritmice (c,d) ale Romanianului.

Testarea acestei supoziții se realizează în următoarele etape de prelucrare :

- transformarea valorilor originale prin logaritmare (indiferent de baza numerică; pentru aplicație s-a utilizat baza $e = 2.718282$);
- analiza grafică a histogramei și diagramei de probabilitate realizată cu valorile logaritmăte;
- testarea analitică a concordanței repartiției valorilor logaritmăte cu modelul repartiției normale.

Analiza comparativă a histogramei și diagramei de probabilitate realizate atât pentru valorile netransformate ale grosimilor Romanianului cât și pentru cele logaritmăte, permit formularea următoarelor observații:

- prin logaritmare, distribuția valorilor grosimii se simetrizează ($\beta_1 = 0,05$) ;
- coliniarizarea punctelor în diagrama de probabilitate realizată cu valorile logaritmăte sugerează o distribuție lognormală a grosimilor;
- se detașează două valori extreme, câte una la ambele extreme (vizibile în diagrama de probabilitate), a căror apartenență la selecție trebuie testată.

Tabelul 1.3. Grosimi ale depozitelor romaniene din Bazinul Dacic.

Nr.	Gr. Rom	Nr.	Gr. Rom	Nr.	Gr. Rom	Nr.	Gr. Rom	Nr.	Gr. Rom	Nr.	Gr. Rom
1	120	18	40	35	90	52	170	69	761	86	360
2	86	19	30	36	190	53	170	70	800	87	301
3	120	20	35	37	100	54	230	71	400	88	200
4	69	21	320	38	160	55	95	72	380	89	200
5	38	22	320	39	180	56	225	73	320	90	452
6	78	23	130	40	300	57	445	74	400	91	392
7	135	24	100	41	320	58	140	75	375	92	233
8	326	25	90	42	300	59	170	76	975	93	1240
9	280	26	100	43	455	60	505	77	305	94	1830
10	90	27	200	44	115	61	792	78	630	95	1426
11	190	28	220	45	420	62	500	79	500	96	1370
12	150	29	200	46	350	63	360	80	1100	97	1364
13	102	30	200	47	486	64	585	81	1000	98	668
14	70	31	200	48	120	65	45	82	2000	99	580
15	200	32	130	49	625	66	220	83	384	100	400
16	47	33	65	50	30	67	680	84	442	101	101
17	250	34	150	51	290	68	1600	85	398		

Datele necesare testării concordanței repartiției valorilor logaritmăte ale grosimii depozitelor romaniene cu modelul repartiției normale, pe baza testului χ^2 sunt: numărul de valori ($n = 101$), media ($m = 5,52$), abaterea standard ($s = 0,95$), numărul intervalelor de grupare ($k = 15$), mărimea intervalului de grupare ($\Delta = 0,3$).

Analog cu aplicația A1 și utilizând datele din **tabelul 1.4**, se obțin, pentru un risc al erorii de genul I de 5%: $\chi^2_{\text{exp}} = 20,85$ și $\chi^2(0,05;12) = 21,03$. Relația dintre χ^2_{exp} și χ^2 indică o repartiție normală a valorilor logaritmăte, adică o repartiție lognormală a valorilor originale (valori netransformate).

Tabelul 1.4. Calculele asociate testului χ^2 pentru analiza distribuției grosimii depozitelor romaniene din Bazinul Dacic .

Nr. interval	Lim. inf. interval	Lim. Sup. Interval	X_{ci}	n_i	np_i	$\frac{(n_i - np_i)^2}{np_i}$
1	3.30	3.60	3.45	3.00	1.23	2.56
2	3.60	3.90	3.75	4.00	2.27	1.31
3	3.90	4.20	4.17	1.00	4.72	2.94
4	4.20	4.50	4.40	7.00	6.40	0.06
5	4.50	4.80	4.67	10.00	8.58	0.23
6	4.80	5.10	4.95	7.00	10.66	1.26
7	5.10	5.40	5.27	15.00	12.25	0.62
8	5.40	5.70	5.52	6.00	12.67	3.51
9	5.70	6.00	5.85	20.00	11.89	5.53
10	6.00	6.30	6.15	9.00	10.18	0.14
11	6.30	6.60	6.44	6.00	7.90	0.46
12	6.60	6.90	6.72	4.00	5.71	0.51
13	6.90	7.20	7.01	3.00	3.70	0.13
14	7.20	7.50	7.27	4.00	2.33	1.20
15	7.50	7.80	7.56	2.00	1.28	0.40

Valorile extreme care sunt vizate pentru testarea apartenenței statistice la selecția de date disponibile sunt : $\ln(\text{grosimea minimă}) = 3,4$ și $\ln(\text{grosimea maximă}) = 7,6$.

Pentru a verifica dacă valorile se încadrează în limitele fluctuațiilor normale se aplică *criteriul intervalului de toleranță*, conform căruia *valorile extreme admisibile* pentru o selecție cu *media (m)* și *abaterea standard (s)* se evaluează cu relațiile :

$$v_{\max} = m + z(\alpha) \cdot s \quad \text{și} \quad v_{\min} = m - z(\alpha) \cdot s$$

în care $z(\alpha)$ este coeficientul abaterii limită care se determină din tabelul funcției $1-2\Phi(z)$, iar $\alpha = 1/(2n)$ (n - numărul de valori ale selecției) (Bomboe, P.1979).

Grosimile logaritmice ale depozitelor romaniene fluctuează în mod normal între limitele : $v_{\min} = 5,52 - (2,80 \cdot 0,95) = 2,86$ și $v_{\max} = 5,52 + (2,80 \cdot 0,95) = 8,18$ în care $\alpha = 0,005$ iar $z(0,005) = 2,80$. Comparând *valorile extreme admisibile* cu *valorile extreme ale selecției*, rezultă că acestea din urmă (adică 3,40 și 7,60) se încadrează în limitele fluctuațiilor normale și deci nu trebuie eliminate.

COMENTARIU

Pentru că repartiția valorilor grosimii depozitelor romaniene în jurul mediei de selecție nu este normală (este lognormală !!!), toate estimările statistice și geostatistice realizate cu aceste valori nu vor fi corecte (vor fi supraestimate !!).

Să prelucrezi valorile în această stare de “anormalitate” este ca și cum ai încerca să te înțelegi cu un om turmentat !

Pentru a obține rezultate corecte (evitarea supraestimărilor), este necesară normalizarea repartiției valorilor înaintea prelucrărilor, ... trezirea lor din beție.

A3. Analiza unei repartiții lognormale și neomogene

Ce distribuție au conductivitățile hidraulice determinate în 113 foraje hidrogeologice (de explorare și captare) executate între Dâmbovița și Ialomița pe o suprafață de aproximativ 200 km² (Fig.1.5) ?

Rezolvare :

Un caz tipic de variabilă cu distribuție eterogenă și lognormală îl constituie *conductivitatea hidraulică* a depozitelor din conurile aluvionare.

Eterogenitatea statistică își găsește explicația în diversitatea condițiilor de formare și a surselor ce participă la constituirea depozitelor aluvionare. Distribuția asimetrică (lognormală) a valorilor conductivității hidraulice în raport cu media de selecție comportă interpretări a căror complexitate crește mult gradul de incertitudine al estimărilor geostatistice.

Analiza distribuției valorilor demarează cu examinarea *histogramei* valorilor (Fig.1.6a) și a *diagramei de probabilitate* (Fig.1.6b).

Cele 113 valori sunt cuprinse în intervalul [6,86 ; 2304,00 m/zi] iar *histograma* indică o puternică asimetrie de stânga. Apare deci necesitatea normalizării distribuției valorilor printr-o transformare numerică a acestora.

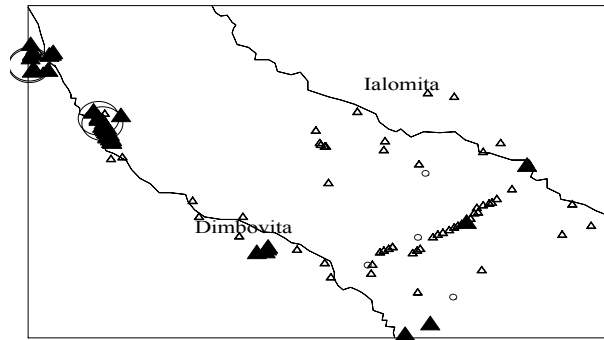


Fig.1.5. Locațiile forajelor hidrogeologice

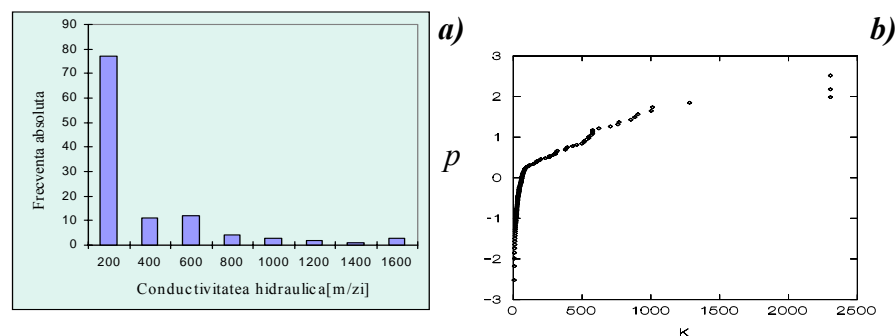


Fig.1.6. Histograma (a) și diagrama de probabilitate (b) pentru valorile conductivităților hidraulice

Diagrama de probabilitate evidențiază asimetria distribuției prin dispunerea punctelor de-a lungul unei curbe. Dacă distribuția valorilor ar fi fost *normală* (simetrică) punctele din diagrama de probabilitate s-ar fi poziționat de-a lungul unei *linii drepte*.

Diagrama de probabilitate permite sesizarea și a unor aspecte suplimentare care servesc analizei *eterogenității* și a *valorilor extreme*:

- alinierea punctelor din diagramă pe *două segmente liniare* sugerează separarea a două grupuri de valori cu caracteristici statistice distincte ;

- se detașează un număr de *opt valori*, câte patru la fiecare extremitate a selecției, care urmează să fie supuse testelor specifice (*criteriul abaterii normate, criteriul limitei de toleranță, criteriul comparării abaterilor etc.*; **tabelul 1.5**).

Eta pa care urmează examinării *histogramei* și *diagramei de probabilitate* este consacrată *normalizării distribuției* selecției de valori care în cazul acesta se realizează prin *logaritmare*. Efectul *logaritmării* este semnificativ, coeficientul de asimetrie al distribuției valorilor *conductivității* reducându-se de la $\beta_3 = 3,07$ (în cazul valorilor originale) la $\beta_3 = 0,36$ (pentru valorile *logaritmăte*).

Tabelul 1.5. Valorile extreme care se elimină din selecție

Nr. Crt.	Valori eliminate		Foraj	Semn conv.pe hartă(Fig.1.5)
	K[m/zi]	Ln(K)		
1	6,86	1.92	9	Cerc mic
2	8,00	2.07	135	Cerc mic
3	8,40	2.12	27	Cerc mic
4	9,00	2.19	105	Cerc mic
5	1281,00	7.15	69	Cerc mare
6	2304,00	7.74	31	Cerc mare
7	2304,00	7.74	33	Cerc mare
8	2304,00	7.74	73	Cerc mare

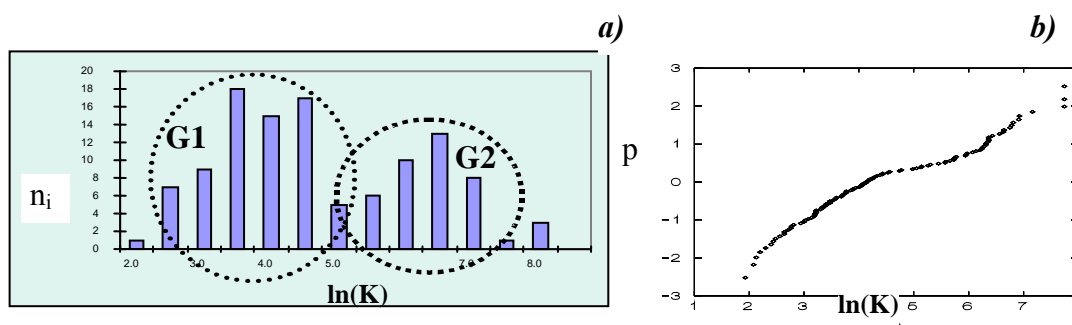


Fig.1.7. Histograma (a) și diagrama de probabilitate (b) ale valorilor *logaritmăte* ale *conductivității hidraulice*.

Analiza morfologiei *histogramei* și a *diagramei de probabilitate* pentru valorile *logaritmăte* (**Fig.1.7a** și **Fig.1.7b**) conduce la următoarele măsuri ce trebuie luate:

- din cele 113 valori de *conductivitate* trebuie eliminate cele *opt valori extreme*, vizibil detașate și în *diagrama de probabilitate* a valorilor *logaritmăte*;
- cele 105 valori rămase după eliminarea valorilor extreme trebuie separate în două grupuri (G1 și G2), grupuri a căror *comparare statistică* se realizează prin *analiza dispersională monofactorială*. Cele două grupuri sunt caracterizate prin elementele sintetizate în **tabelul 1.6**. Ele se diferențiază net doar prin valorile *mediilor*, având practic aceeași *dispersie* a valorilor.

Tabelul 1.6. Caracteristici statistice ale grupurilor G1 și G2

Grup	Număr de valori	Valoarea medie ln(K)/K	Coef. de asimetrie	Dispersia
G1	65	3,54/ 34,47	-0,24	0,36
G2	40	5,98/ 395,72	-0,39	0,37

Valorile din grupul *G1* sunt mai mici și au fost determinate în interfluviul Dâmbovița-Ialomița iar valorile grupului *G2*, mai mari, au fost determinate în foraje plasate de-a lungul văii Dâmbovița (**Fig.1.5**).

Separarea statistică a celor două grupuri corespunde în acest caz cu condițiile distincte în care s-au depus formațiunile a căror conductivitate hidraulică se studiază: în vecinătatea talvegului văilor, unde sunt formațiuni mai grosiere și bine spălate, sunt depozite cu conductivități mari (*G2*) iar în interfluviu sunt depozite cu granulație mai fină și uneori cu matrice argiloasă, deci cu conductivitate hidraulică mai mică (*G1*).

Considerând că diferențierea statistică a celor două grupuri de valori este determinată de *un singur factor* sau de *mai mulți factori cu acțiune convergentă*, instrumentul adecvat testării statistice a acestei ipoteze este *analiza dispersională monofactorială*.

La o eroare de genul I cu $\alpha = 5\%$, raportul dintre factorul Fischer experimental ($F_{exp.} = 93,67$) și cel critic ($F_{cr} = 3,93$) este $F_{exp}/F_{cr} = 23,83$. Aceasta justifică din punct de vedere statistic separarea celor două grupuri de valori. În continuarea prelucrărilor geostatistice, cele două grupuri de valori logaritmice ale conductivității hidraulice vor fi considerate ca două selecții de valori distincte, pentru care prelucrările se vor face separat (**Tabelul 1.7** în care valorile conductivităților hidraulice din paranteze sunt logaritmice).

Tabelul 1.7. Amplitudinea conductivității hidraulice din *G1* și *G2*.

Grup	Valoarea minimă	Valoare maximă	Nr.valori
<i>G1</i>	10 (2,3)	100 (4,6)	65
<i>G2</i>	100 (4,6)	1011 (6,91)	40

Este important de reținut că :

- *normalizarea* distribuției valorilor variabilelor studiate (prin logaritmă în cazul aplicației) este necesară pentru a elimina riscul supraestimării/subestimării evaluărilor geostatistice;
- *separarea selecțiilor neomogene în grupuri de valori omogene* (în cazul aplicației grupurile *G1* și *G2*) permite identificarea legităților de variație spațială specifice domeniilor în care procesul depunerii depozitelor permeabile a fost controlat de factori diferiți.

Finalizarea analizei variabilității globale, când ea nu este urmată de o evaluare geostatistică (spațială), o constituie *valoarea medie* a parametrului studiat asociată cu un *interval de încredere* corespunzător *erorii de genul I* alese.

Pentru conductivitatea hidraulică a depozitelor din zona cercetată, analiza variabilității globale conduce la concluzia că depozitele conului aluvionar se diferențiază net în două zone:

- zona talvegului râului Dâmbovița (*G2*) unde valoarea medie a conductivității este:

$$K_2 = 394 \text{ m/zi} \pm 1 \text{ m/zi pentru } \alpha = 5\%;$$

- zona interfluviului Dâmbovița-Ialomița (*G1*) unde valoarea medie a conductivității este:

$$K_1 = 34 \text{ m/zi} \pm 2 \text{ m/zi pentru } \alpha = 5\% .$$

Neglijarea *distribuției lognormale* a selecției de date și *neomogenitatea ei statistică* ar fi condus la *subestimări* ale conductivității hidraulice în zona talvegului

Dâmboviței (G_2) și la *supraestimări* în interfluviul Dâmbovița-Ialomița, cu efecte importante asupra studiului curgerii apelor subterane.

Pentru evaluarea valorii medii a conductivității în zona talvegului Dâmboviței se calculează *media aritmetică ponderată* a celor $n_2 = 40$ valori logaritmice ale conductivității hidraulice din grupul G_2 și se obține $\ln(K_2) = 5,977$ care prin antilogaritmare conduce la $K_2 = 394,35$ m/zi, rotunjit la $K_2 = 394$ m/zi prin trunchare.

Intervalul de încredere al mediei, pentru $\alpha = 5\%$ se calculează tot prin intermediul celor $n_2 = 40$ valori logaritmice ale conductivităților hidraulice obținându-se $\ln(\varepsilon) = 0,194$ care prin antilogaritmare conduce la $\varepsilon = 1,2$ m/zi.

Valorile pentru zona interfluviului Dâmbovița-Ialomița (G_1) (valoarea medie a conductivității hidraulice și a erorii de estimare) s-au calculat prin prelucrarea celor $n_1 = 65$ de valori logaritmice ale conductivității hidraulice din grupul G_1 , după metodologia explicitată pentru G_2 .

COMENTARIU

Conductivitățile hidraulice determinate în 113 foraje hidrogeologice executate între Dâmbovița și Ialomița (Fig.1.5) formează o selecție de date neomogenă și cu repartiție lognormală.

Cum reacționăm în astfel de situații?

Creștem prețul prelucrării deoarece efortul necesar pregătirii datelor pentru realizarea unor evaluări geostatistice corecte crește exponențial. Prezența neomogenităților statistice dă serioase dureri de cap geologilor, deoarece:

- *dacă neglijează aceste neomogenități se pierde în totalitate controlul erorilor făcute prin estimări (nu mai știi dacă faci supraestimări sau subestimări!!);*
- *dacă se dorește eliminarea efectului lor, apar dificultăți procedurale și conceptuale greu de depășit.*

Eliminarea efectului neomogenităților statistice presupune cunoașterea în profunzime a fenomenului studiat (în cazul aplicației - procesul de formare a depozitelor permeabile) și particularitățile metodei statistice (în cazul aplicației - analiza dispersională monofactorială).

*Cheia rezolvării corecte a problemei neomogenității selecțiilor de date este separarea corectă a grupurilor de date omogene (în cazul aplicației: G_1 și G_2). Din punct de vedere statistic separarea în grupuri omogene nu este unică și numai cunoașterea corectă a procesului natural studiat ne permite să alegem soluția adecvată. Mai laborioasă și mai complexă decât analiza modului de distribuție a valorilor sau analiza valorilor extreme, cea de-a treia componentă a **analizei variabilității globale**, analiza eterogenității selecțiilor de date geologice, este posibilă numai pe baza cunoașterii profunde a proceselor geologice.*

Fără cunoașterea condițiilor geologice în care se desfășoară fenomenul studiat, analiza eterogenității selecțiilor de date este un simplu exercițiu statistic, fără semnificație geologică !

1.2. NORMALIZAREA REPARTIȚIEI SELECȚIILOR DE DATE

Este puțin probabil ca toți să percepem realitatea la fel. Am convingerea că de fapt totul se reglează prin limbaj și chiar dacă spunem toți că *vedem un măr roșu* (al lui Machintosh, de exemplu), senzațiile fiecăruia diferă; altfel spus, *fiecare cu mărul lui roșu*. Și dacă punem zece pictori să deseneze același măr roșu vom avea zece viziuni diferite ale mărului roșu.

Este dificil să privim toți realitatea prin aceeași *lentilă*.

Fiecare cu dioptriile lui.

Pentru o scurtă perioadă, cea a *prelucrărilor geostatistice*, vă propun să acceptați acest lucru. Este în interesul unei *estimări obiective* a realității, estimare realizată prin filtrul instrumentelor matematice. *Lentila* prin care vom privi *realitatea* ne va face s-o vedem totdeauna *normală*. Este o *lentilă interactivă*, este cunoscută sub denumirea de *normalizarea* realității și la orice *anormalitate* a acesteia distorsionează semnalul plasându-l în coordonatele unui model unic, cel al *distribuției normale (gaussiene)*.

De ce este nevoie să acceptăm această lentilă ?

Cițiți mai departe și sper că veți fi cel puțin curioși să încercați o senzație nouă, aceea de *normalitate generalizată*.

Metodele geostatistice (topo-probabiliste) sunt puse la punct pentru prelucrarea selecțiilor de date cu distribuție *normală (gaussiană)*. Această premiză nu exclude utilizarea acestor metode și pentru variabilele cu altfel de distribuții. Dacă valorile variabilelor prelucrate (v_i) se abat de la repartiția normală, aplicarea corectă a *metodelor geostatistice (topo-probabiliste)* necesită o transformare a datelor originale ($T(v_i)$) care să conducă la valori cu distribuție normală (t_i):

$$t_i = T(v_i) \quad (1.1)$$

Valorile transformate vor fi prelucrate cu metodologia specifică modelelor topo-probabiliste. La finalul prelucrărilor pentru revenirea în câmpul valorilor originale se realizează transformarea inversă (T^{-1}) celei prin care datele originale au fost transformate în vederea prelucrării.

Normalizarea distribuției diferitelor variabile poate fi realizată în *câmpul valorilor normale* și este cunoscută sub numele de *normalizare redusă*. *Normalizarea redusă* a distribuțiilor în *câmpul valorilor normale* conduce la o *variabilă normată* (u_i) cu *repartiție normală*, cu *media zero* ($m = 0$) și *dispersia unitară* ($s^2 = 1$). Pentru *normarea* valorilor v_i se utilizează relația :

$$u_i = \frac{v_i - m}{s} \quad (1.2)$$

în care u_i este valoarea normată, m și s sunt *media de selecție*, respectiv *abaterea standard de selecție* a valorilor netransformate (v_i). Normalizarea valorilor poate fi realizată și în câmpul valorilor originale fără normarea acestora și este cunoscută sub denumirea de *normalizare generalizată*.

Normalizarea distribuției valorilor se bazează pe *probabilitatea de apariție* (p_i) a fiecărei valori măsurate (v_i) ($i = 1 \dots n$, n - numărul total de valori măsurate).

Probabilitățile de apariție a valorilor măsurate (p_i), în cazul variabilelor geologice, de cele mai multe ori depind de distribuția în spațiu a punctelor în care se face determinarea lor:

a) dacă punctele de observație sunt distribuite uniform pe suprafața cercetată această probabilitate se aproximează prin relația:

$$p_1 = p_2 = \dots = p_n = \frac{1}{n} \quad (1.3)$$

b) dacă punctele de observație sunt distribuite neuniform pe suprafața cercetată, probabilitățile pot fi estimate prin diferite tehnici (declustering celular, declustering poligonal etc., Scrădeanu, D., 1996).

În ambele cazuri (a și b), trebuie respectată condiția :

$$\sum_{i=1}^n p_i = 1 \quad (1.4)$$

Calculul probabilităților pentru valorile extreme (valori maxime și minime) ale selecțiilor de date trebuie abordat în mod diferențiat. Pentru situația unor volume reduse de date trebuie luată în considerare o probabilitate diferită de zero pentru valori mai mici ca valoarea minimă și mai mari decât cea maximă.

O soluție simplistă pentru calculul probabilităților valorilor extreme nedeterminate este egalarea sumei probabilităților cu o valoare mai mică decât unitatea (ex.: $n/(n+1)$), soluție sensibilă însă la variația numărului de probe disponibile.

Normalizarea valorilor măsurate se percepe cel mai comod pe baza unei reprezentări grafice (Fig.1.8).

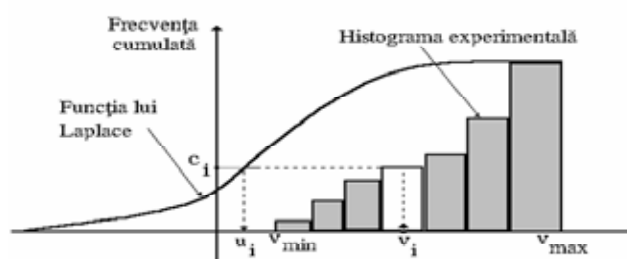


Fig.1.8.Principiul normalizării valorilor.

Pentru această operație sunt necesare două curbe de frecvențe cumulate:

- curba frecvențelor cumulate a valorilor măsurate (v_i), adică histograma experimentală cumulată;
- curba frecvențelor cumulate ale repartiției normale (funcția lui Laplace ($\Phi(u)$)).

Pentru normalizarea valorii v_i se duce prin valoarea măsurată (v_i) o paralelă la axa frecvențelor până ce intersectează curba frecvențelor cumulate a valorilor măsurate. Din punctul de intersecție se duce o paralelă la abscisă până intersectează funcția lui Laplace iar de aici o paralelă la axa frecvențelor obținându-se valoarea normalizată căutată (u_i).

Echivalența analitică a acestei operațiuni grafice este:

$$u_i = G^{-1}(c_i) \quad (1.5)$$

în care

G^{-1} - inversa integralei lui Gauss;

c_i - probabilitatea cumulată corespunzătoare valorii v_i .

Normalizarea generalizată a distribuției datelor în *câmpul valorilor reale* este o operațiune similară cu cea a normalizării în câmpul valorilor normate care presupune înlocuirea valorilor funcției lui Laplace cu o serie de valori cu repartiție normală de medie și dispersie cunoscute. Media și dispersia se aleg în funcție de valorile a căror distribuție se normalizează.

Prin această operațiune se poate transforma distribuția oricărui set de valori în raport cu o *distribuție de referință* (reprezentată printr-un alt set de valori) fără a se cunoaște modelul analitic al acestei distribuții.

A4. Normalizare redusă

Să se normalizeze, în spațiul valorilor normate, cele 124 de valori ale grosimii depozitelor daciene (determinate în 124 de foraje; **Fig.1.9**).

Rezolvare :

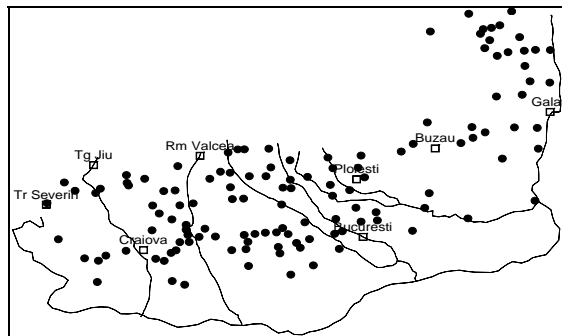


Fig.1.9. Distribuția celor 124 de foraje în care a fost măsurată grosimea Dacianului.

Din cauza *asimetriei de stânga*, prelucrarea acestor valori cu modelele topo-probabiliste de tipul *kriging*-ului va conduce la *supraestimări* în evaluarea distribuției spațiale a grosimii Dacianului din Bazinul Dacic

Pentru realizarea *normalizării* distribuției grosimilor se utilizează inversa integralei lui Gauss (G^{-1}) obținută prin aproximarea numerică Kennedy and Gentle (1980).

Un program comod de utilizat este disponibil în biblioteca GSLIB (*nscore.f*). Trebuie precizat că programul:

- utilizează ca date de intrare: un *fișier cu date* de tip Geo-EAS (**Fig.1.11a**) și un *fișier cu parametrii* de prelucrare (**Fig.1.11b**);
- scrie rezultatele în două fișiere: un *fișier cu toate datele* (**Fig.1.12a**) și un *fișier cu valorile măsurate și transformate* (**Fig.1.12b**) utilizat la transformarea inversă, de la valorile reduse cu distribuție normală (u) la cele măsurate (v).

Considerăm, într-o primă aproximare, că probabilitățile de apariție pentru fiecare din cele 124 de grosimi determinate în cele 124 de foraje (**tabelul 1.8**) sunt egale între ele și egale cu $p = 1/124$.

Histograma și curba cumulativă a frecvențelor (**Fig.1.10**) pentru grosimile Dacianului indică o repartiție asimetrică cu un *coeficient de asimetrie* $\beta_1 = 1.09$.

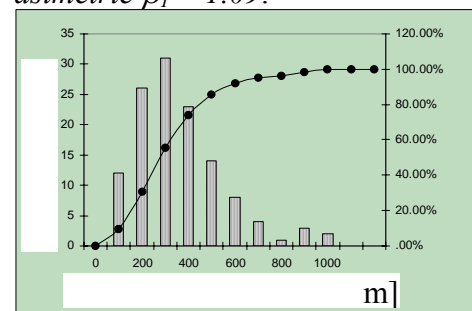


Fig.1.10. Histograma și curba cumulativă a grosimii depozitelor daciene

În cazul aplicației, fișierul cu date de tip Geo-EAS conține 4 coloane pe care sunt inserate: *numărul curent al forajului*, *abscisa forajului*, *ordonata forajului* și *valoarea grosimii Dacianului* (**Fig.1.11a**)

Fișierul cu parametrii de transformare (**Fig.1.11b**), în cazul aplicației, conține:

- *numele fișierului* cu date (ex.: *dacian.dat*);
- *numărul coloanelor* pe care se află:
 - *valoarea variabilei*, adică 4, deoarece grosimea Dacianului se află în fișierul de date pe coloana a patra;
 - *ponderile* acordate acestor valori, adică 0 (zero), deoarece am plecat de la ipoteza că fiecare valoare are probabilitatea 1/124 și în tabelul cu date nu există o coloană cu valorile acestor ponderi.
- *valoarea maximă* și cea *minimă* luate în considerare pentru transformare:
 - grosimea minimă = 15 m
 - grosimea maximă = 928 m

<p>GROSIMILE DACIANULUI</p> <p>4 a)</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 10%;">Nr.crt.</td> <td style="width: 10%;">_</td> <td style="width: 10%;">G16.9</td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> </tr> <tr> <td>x</td> <td>m</td> <td>G16.9</td> <td></td> <td></td> <td></td> </tr> <tr> <td>y</td> <td>m</td> <td>G16.9</td> <td></td> <td></td> <td></td> </tr> <tr> <td>grosime</td> <td>m</td> <td>G16.9</td> <td></td> <td></td> <td></td> </tr> <tr> <td>1</td> <td>305.20</td> <td>96.80</td> <td>429.00</td> <td></td> <td></td> </tr> <tr> <td>2</td> <td>364.77</td> <td>249.06</td> <td>430.00</td> <td></td> <td></td> </tr> <tr> <td>3</td> <td>335.38</td> <td>247.32</td> <td>905.00</td> <td></td> <td></td> </tr> <tr> <td colspan="6" style="border-top: 1px dashed black;"></td> </tr> <tr> <td>124</td> <td>46.00</td> <td>100.44</td> <td>191.00</td> <td></td> <td></td> </tr> </table>	Nr.crt.	_	G16.9				x	m	G16.9				y	m	G16.9				grosime	m	G16.9				1	305.20	96.80	429.00			2	364.77	249.06	430.00			3	335.38	247.32	905.00									124	46.00	100.44	191.00			<p style="text-align: right;">b)</p> <p>PARAMETRII DE TRANSFORMARE</p> <p>dacian.dat</p> <p>4 0</p> <p>15 928</p> <p>ndacian.dat</p> <p>ndacian.trn</p>
Nr.crt.	_	G16.9																																																					
x	m	G16.9																																																					
y	m	G16.9																																																					
grosime	m	G16.9																																																					
1	305.20	96.80	429.00																																																				
2	364.77	249.06	430.00																																																				
3	335.38	247.32	905.00																																																				
124	46.00	100.44	191.00																																																				

Fig.1.11. Structura fișierelor de intrare pentru programul de transformare: **a)** fișierul de date; **b)** fișierul cu parametrii de transformare.

<p>GROSIMILE DACIANULUI</p> <p>5 a)</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 10%;">Nr.crt.</td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> </tr> <tr> <td>x</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>y</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>grosime</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>grosime transformată(normalizată)</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>1</td> <td>305.20</td> <td>96.80</td> <td>429.00</td> <td>0.74</td> <td></td> </tr> <tr> <td>2</td> <td>364.77</td> <td>249.06</td> <td>430.00</td> <td>0.77</td> <td></td> </tr> <tr> <td>3</td> <td>335.38</td> <td>247.32</td> <td>905.00</td> <td>2.25</td> <td></td> </tr> <tr> <td colspan="6" style="border-top: 1px dashed black;"></td> </tr> <tr> <td>127</td> <td>46.00</td> <td>100.44</td> <td>191.00</td> <td>-0.56</td> <td></td> </tr> </table>	Nr.crt.						x						y						grosime						grosime transformată(normalizată)						1	305.20	96.80	429.00	0.74		2	364.77	249.06	430.00	0.77		3	335.38	247.32	905.00	2.25								127	46.00	100.44	191.00	-0.56		<p style="text-align: right;">b)</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> </tr> <tr> <td></td> <td></td> <td></td> <td>429.00</td> <td>0.74</td> <td></td> </tr> <tr> <td></td> <td></td> <td></td> <td>430.00</td> <td>0.77</td> <td></td> </tr> <tr> <td></td> <td></td> <td></td> <td>905.00</td> <td>2.25</td> <td></td> </tr> <tr> <td colspan="6" style="border-top: 1px dashed black;"></td> </tr> <tr> <td></td> <td></td> <td></td> <td>191.00</td> <td>-0.56</td> <td></td> </tr> </table>										429.00	0.74					430.00	0.77					905.00	2.25											191.00	-0.56	
Nr.crt.																																																																																																	
x																																																																																																	
y																																																																																																	
grosime																																																																																																	
grosime transformată(normalizată)																																																																																																	
1	305.20	96.80	429.00	0.74																																																																																													
2	364.77	249.06	430.00	0.77																																																																																													
3	335.38	247.32	905.00	2.25																																																																																													
127	46.00	100.44	191.00	-0.56																																																																																													
			429.00	0.74																																																																																													
			430.00	0.77																																																																																													
			905.00	2.25																																																																																													
			191.00	-0.56																																																																																													

Fig.1.12. Fișierele cu rezultatele transformării (normalizării) **a)** fișierul cu rezultatele normalizării (*ndacian.dat*); **b)** fișierul cu tabelul utilizat la transformarea inversă (*ndacian.trn*).

- *numele fișierului* cu rezultatele transformării (ex.: *ndacian.dat*); acest fișier, pe lângă datele din fișierul de intrare (*dacian.dat*) conține și coloana valorilor transformate (**Fig.1.12a**);
- *numele fișierului* cu tabelul de valori utilizat la transformarea inversă care se realizează după prelucrarea valorilor transformate: *ndacian.trn* (**Fig.1.12b**).

Valorile transformate (normalizate), cele din coloana a 5-a din fișierul *ndacian.dat*, au o repartiție perfect simetrică (**Fig.1.13**), media zero și dispersia unitară. Aceste valori se utilizează în etapa de interpolare cu ajutorul *kriging*-ului.

După realizarea operațiunii de interpolare, pe baza valorilor transformate, rezultatele finale se proiectează în câmpul valorilor reale prin intermediul corelației dintre valorile măsurate și cele transformate (**Fig.1.14**), conținute în fișierul *ndacian.trn* sau pe coloanele 4 și 5 din fișierul *ndacian.dat*.

Aspectul grafic al corelației dintre cele două șiruri de valori arată gradul de abatere al valorilor măsurate (grosimile Dacianului) de la repartiția normală. Dacă valorile măsurate ale grosimilor ar fi avut o repartiție normală punctele din **Fig.1.14**. s-ar fi aliniat după o dreaptă.

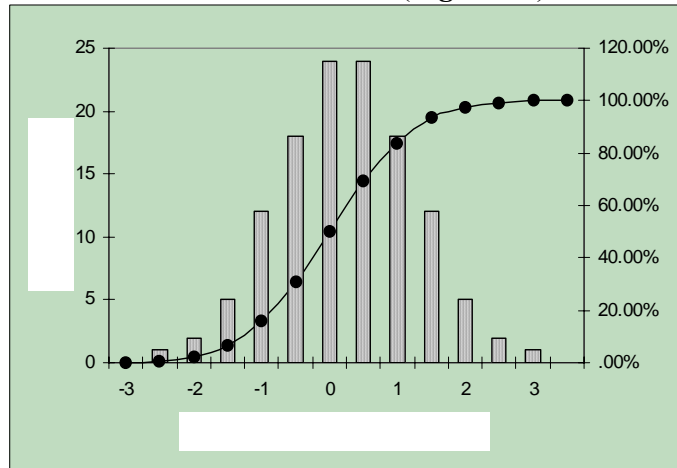


Fig.1.13. Histograma și curba frecvențelor cumulate pentru grosimile transformate ale Dacianului, obținute prin normalizare redusă

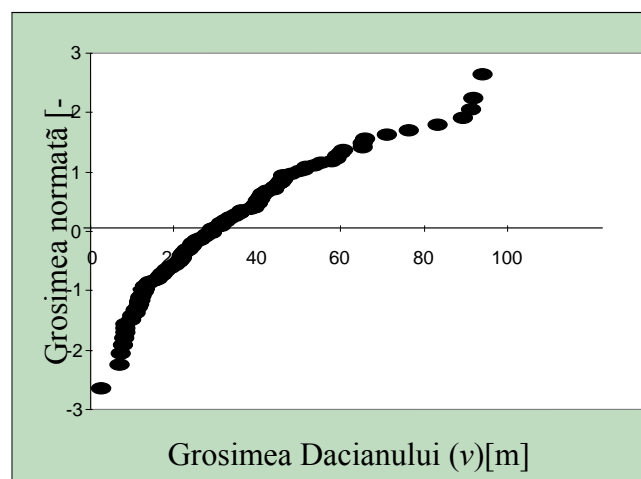


Fig.1.14. Corelația dintre grosimile măsurate (*v*) și cele transformate (*u*, normale și normalizate).

COMENTARIU

Normalizarea este trezirea din beție a datelor.

Normalizarea selecțiilor de date

*Necesitatea normalizării distribuției datelor rezultă din analiza morfologiei **histogramei și diagramei de probabilitate**, aceleași instrumente care sunt utilizate și pentru **verificarea** rapidă a eficienței operațiunii de normalizare.*

Dacă după normalizare histograma valorilor este simetrică iar în diagrama de probabilitate punctele sunt coliniare operațiunea a reușit.

*Cu valorile transformate, **trezite din beție**, se poate lucra în mod **normal**.*

După estimările geostatistice, pentru a respecta opțiunile și libertățile, se poate reveni la starea anterioară prin procedeul foarte bine cunoscut.

A5. Normalizare generalizată

Să se normalizeze distribuția grosimilor depozitelor daciene (v), determinate în 124 foraje (**Fig.1.9** și **tabelul 1.8**) pentru domeniul valoric al selecției disponibile.

Rezolvare:

Operarea cu valori adimensionale, cuprinse în domeniul valoric $[-3, 3]$, obținute prin *normalizare redusă*, chiar dacă este confortabilă din punct de vedere numeric, poate crea dificultăți de interpretare pe parcursul operării. Pentru reducerea acestor dificultăți formale mulți specialiști preferă utilizarea *normalizării generalizate* care în urma transformării proiectează valorile originale tot în domeniul valoric în care au fost determinate dar sub forma unei distribuții normale

Normalizarea generalizată implică două etape de prelucrare:

- 1) generarea unui set de valori cu repartiție normală;
- 2) *normalizarea* valorilor originale, utilizând ca repartiție de referință repartiția setului de date generat.

Generarea setului de valori cu repartiție normală se poate realiza prin proiectare lineară a setului de valori obținute prin *normalizare redusă* în aplicația A3 (cuprins în intervalul $[-3,3]$), pe domeniul valoric al celor 124 de grosimi măsurate [$v_{\min} = 15\text{m}$, $v_{\max} = 928\text{m}$], prin relația:

$$v_i = v_{\min} + \frac{v_{\max} - v_{\min}}{u_{\max} - u_{\min}} \cdot (u_i - u_{\min})$$

în care

u_{\min} , u_i , u_{\max} - valori normate cu repartiție normală (**Fig.1.15a**);

v_{\min} , v_{\max} - valoarea minimă, respectiv valoarea maximă măsurate;

v_i - valorile cu repartiție normală utilizate ca repartiție de referință (**Fig.1.15b**)

În cazul aplicației valorile utilizate pentru calcule sunt:

$$u_{\min} = -3, \quad u_{\max} = 3, \quad v_{\min} = 15\text{m}, \quad v_{\max} = 928\text{m}$$

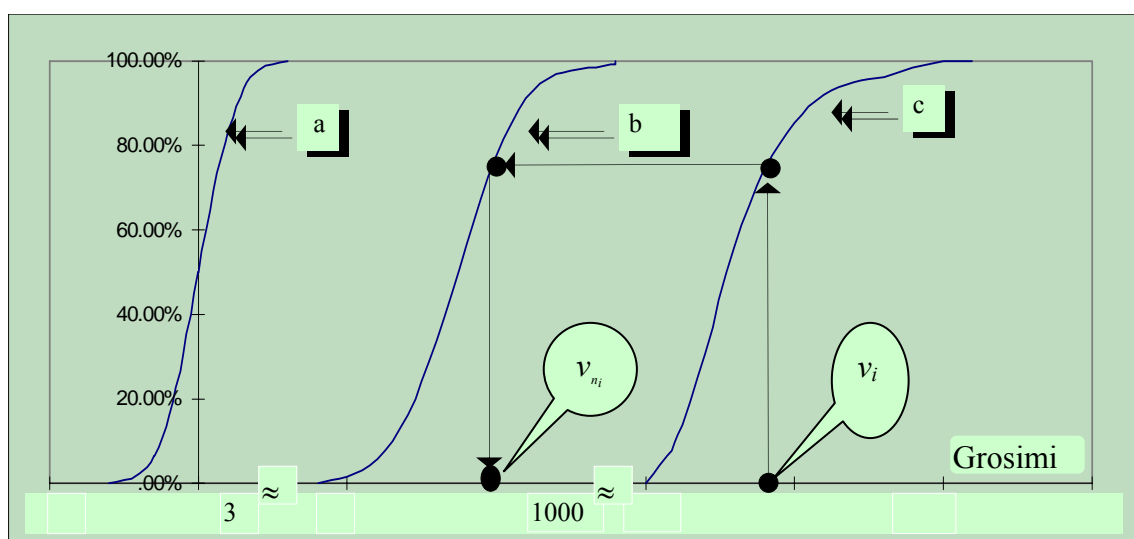


Fig.1.15. Curbele frecvențelor cumulate necesare normalizării generalizate:
a) curba frecvențelor cumulate a grosimilor normate cu repartiție normală;
b) curba de referință a frecvențelor cumulate cu repartiție normală;
c) curba experimentală a frecvențelor cumulate pentru grosimile măsurate.

Normalizarea valorilor măsurate în câmpul valorilor reale al grosimilor, se realizează grafic utilizând curba de referință a frecvențelor cumulate (**Fig.1.15b**) obținută în etapa anterioară și curba experimentală a frecvențelor cumulate pentru grosimile măsurate (**Fig.1.15c**).

Correspondentul analitic al procedurii grafic indicat prin săgeată în **Fig.1.15** este relația:

$$v_{n_i} = F^{-1}(c_i)$$

în care

v_{n_i} - valoarea transformată în concordanță cu repartiția normală de referință;

c_i - frecvența cumulată;

F^{-1} - inversa funcției experimentale de frecvență.

Valorile obținute prin această transformare (v_{n_i}) au repartiție normală și sunt cuprinse în intervalul [15; 928] (**Fig.1.16**).

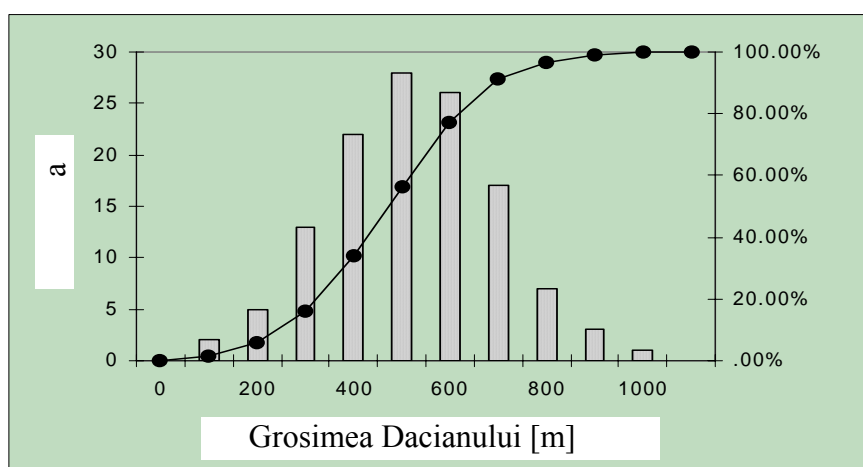


Fig.1.16. Histograma și curba frecvențelor cumulate pentru grosimile Dacianului obținute prin normalizare generalizată.

COMENTARIU

Orice selecție de valori care **nu are distribuție normală**, înaintea oricăror evaluări statistice, trebuie **normalizată**. Ignorarea acestei etape de prelucrare conduce cu siguranță la supraestimări sau subestimări.

Este preferabil să **normalizăm o selecție de valori cu distribuție normală** (gaussiană) decât să nu normalizăm o selecție cu distribuție negaussiană.

Normalizarea unei selecții de valori cu distribuție normală nu aduce nici o modificare a valorilor prin transformările aplicate (normalizare redusă/normalizare generalizată). Această variantă de prelucrare poate fi o alternativă pentru **testarea normalității** distribuției unei selecții de valori.

1.3. ANALIZA VARIABILITĂȚII SPAȚIALE

Analiza variabilității spațiale integrează în prelucrare, pe lângă valorile caracteristicilor geologice (v_i) care au constituit obiectul prelucrării în cadrul *analizei variabilității globale*, o a doua categorie de date: *coordonatele spațiale* (x_i, y_i, z_i) ale punctelor în care au fost determinate valorile acelor caracteristici.

Cele n valori v_i ($i = 1 \dots n$) pe care până acum le-am ținut într-un sac pe care l-am scuturat până ce valorile au fost *normalizate* ($t_i = T(v_i)$) sunt acum împrăștiate pe suprafața de unde au fost colectate pentru a obține elementele necesare realizării desenului ce redă forma obiectelor din hărțile și secțiunile geologice.

Variabilitatea spațială, de mare complexitate pentru caracteristicile geologice, este obiectul unor metode de analiză și sinteză foarte laborioase. Eficiența acestor metode, proporțional cu gradul lor de sinteză, este determinată de experiența celor care le aplică și de parcurgerea într-o succesiune strictă a următoarelor etape:

- *reprezentarea grafică* a datelor, utilizată pentru formarea unei imagini generale asupra distribuției valorilor variabilei în spațiul cercetat;
- *analiza parametrică* a datelor ce sintetizează în trei tipuri de *funcții de distanță* caracteristicile variabilității spațiale atât pentru o *singură variabilă* (*covarianța, corelograma și variograma*) cât și pentru o *pereche de variabile* (*intercovarianța, intercorelograma și intervariograma*) probate în aceleași puncte de observație;
- *analiza staționarității* caracteristicilor geologice studiate, primul pas dificil al analizei variabilității spațiale, care pune la încercare experiența cercetătorului privind circumstanțele acceptării unor aproximări și a consecințelor acestor aproximări asupra rezultatelor finale ale prelucrărilor geostatistice;
- *analiza variografică*, ultima și cea mai dificilă etapă a analizei variabilității spațiale în care cuantificarea variabilității spațiale se face sub forma *cele mai probabile legi de variație spațială* a caracteristicii studiate.

1.3.1. Reprezentarea grafică

Proprietăți precum *localizarea valorilor extreme* (minime sau maxime), *tendința de evoluție regională*, *gradul de continuitate*, sunt de mare interes pentru studiul proceselor geologice.

Bazată pe un număr minim de instrumente și prelucrări, *reprezentarea grafică* are ca obiectiv sintetizarea caracteristicilor topologice ale datelor trecute deja prin filtrul analizei *variabilității globale univariate* (*tip de repartiție, valori extreme, dispersie*) și *multivariate* (*analiză discriminant, analiză factorială, analiză corelatorie și spectrală* etc.; D.Scrădeanu, 1995).

Ca și *histograma*, pentru *tipul de repartiție*, sau *dreapta de regresie*, pentru *corelația* dintre două variabile, cele mai eficiente instrumente pentru descrierea variabilității spațiale sunt cele grafice. Principalele caracteristici structurale ale datelor primare se exprimă în mod curent prin: *hărți punctuale, hărți simbolice și indicatoare, diagrame de continuitate și variabilitate*.

Harta punctuală se realizează prin simpla dispunere într-un sistem de coordonate a punctelor de observație lângă care se înscriu sau nu, în funcție de densitatea punctelor de observație, valorile variabilei studiate. Harta punctuală se realizează în prima etapă a studiului caracteristicilor spațiale, ea fiind utilizată pentru:

- identificarea erorilor în amplasarea punctelor de observație;
- calculul densității punctelor de observație;
- localizarea valorilor extreme, determinate fie de erori de măsură, fie de anomalii locale, care solicită un interes special (ex.: prezența unor *pepite*).

Harta punctuală este utilă pentru situația în care *numărul de puncte de observație este redus*; în caz contrar se apelează la *harta simbolică*.

Harta simbolică presupune o primă filtrare a datelor primare prin reducerea variabilității spațiale și se realizează în cazul unui număr mare de puncte de observație, număr care face inexpresivă *harta punctuală* prin suprapunerea punctelor de observație sau a etichetelor atașate acestora.

Suprafața pe care se realizează harta simbolică este acoperită cu o *rețea rectangulară/pătratică* în celulele căreia se calculează valoarea medie a caracteristicii studiate. *Numărul de simboluri* (alfanumerice, tonuri de gri sau culori) utilizate se stabilește în funcție de gradul de detaliere necesar. Când se utilizează numai două simboluri *harta simbolică* poartă denumirea de *hartă indicatoare*.

Dimensiunile celulelor rețelei rectangulare/pătratice, procedeul de calcul al valorii medii pentru fiecare celulă și numărul de simboluri influențează în mod determinant aspectul hărții simbolice.

Hărțile simbolice oferă o imagine simplificată a distribuției spațiale a valorilor caracteristicii studiate și permite sesizarea tendințelor și localizarea zonelor cu valori maxime și minime.

Diagrama de continuitate este reprezentarea grafică prin care se face o primă evaluare a gradului de precizie cu care se va putea calcula distribuția spațială a unei variabile, probată prin intermediul unei rețele de observație cu o geometrie oarecare. Buna continuitate a variabilei (adică o variație lentă a variabilei de la un punct la altul) conduce la o precizie crescută în evaluarea distribuției spațiale. Diagrama de continuitate sintetizează într-o reprezentare rectangulară binară, similaritatea valorilor măsurate în puncte vecine.

Este ușor de intuit că similaritatea a două valori vecine depinde de: distanța care separă punctele în care au fost determinate valorile;

orientarea spațială a dreptei care unește aceste puncte.

Pentru a construi *diagrama de continuitate* se utilizează un *vector de poziție* care este caracterizat prin:

- *modul*, numeric egal cu lungimea segmentului care unește cele două puncte ($|h_{ij}^p|$ pentru punctele P_i și P_j din **Fig. 1.17**);
- *orientarea* acestui segment (θ din **Fig. 1.17**), măsurată prin unghiul dintre direcția axei abscisei și a segmentului care unește cele două puncte.

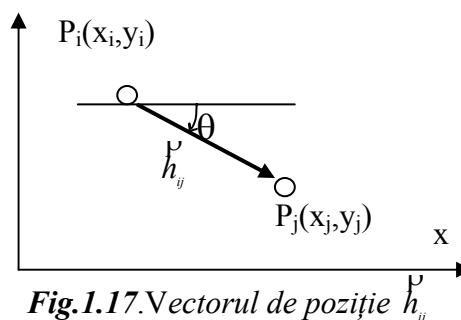


Fig. 1.17. Vectorul de poziție h_{ij}^p

Vectorul de poziție ales este utilizat pentru a identifica toate perechile de puncte din spațiul cercetat care se află la distanța și orientarea aleasă. Perechile de valori $(v_i(P), v_j(P+h_{ij}^P))$ se reprezintă prin puncte într-un sistem de referință rectangular (**Fig.1.18**).

Distribuția punctelor din diagrama de continuitate exprimă gradul de continuitate al variabilei studiate:

- *gruparea* “strânsă” a punctelor în jurul bisectoarei unghiului dintre axele sistemului de referință indică o bună continuitate a variabilei pentru direcția și distanța aleasă; la limită, când distanța pentru care se întocmește diagrama de continuitate este *zero*, toate punctele se află plasate pe bisectoarea unghiului dintre cele două axe ale sistemului de referință deoarece, pentru orice punct, valoarea din acel punct este egală cu ea însăși;
- *dispersarea* punctelor în spațiul dintre cele două axe ale sistemului rectangular de referință indică o slabă continuitate; cu cât norul de puncte este mai difuz, cu atât similaritatea valorilor este mai mică și deci continuitatea mai slabă.

Diagramele de continuitate sunt afectate în mod semnificativ de *valorile extreme* ale variabilei studiate, de *modulul și direcția* vectorului de poziție (h_{ij}^P). Pentru etapa evaluării *modelului de structură spațială (variograma)*, diagramele de continuitate sunt singurele instrumente care permit identificarea valorilor *nodale* (*valori nodale* - valori care sunt determinante în stabilirea legii de variație spațială a unei caracteristici regionalizate) ale structurilor spațiale. Numai *diagramele de continuitate* permit eliminarea *valorilor extreme*, care nu se încadrează în modelul structural fiind “*accidente structurale*”.

Compararea diagramelor de continuitate pe diferite direcții permite identificarea *anizotropiei structurilor spațiale*. Structurile *anizotrope* sunt caracterizate de diagrame de continuitate diferite pe direcții diferite de calcul. Identitatea diagramelor de continuitate calculate pe orice direcție indică *izotropia* structurii.

Diagrama de variabilitate exprimă corelația dintre *valoarea medie* (m) a unei caracteristici într-o anumită zonă și *eroarea* cu care ea poate fi estimată. *Eroarea* este calculată pe baza *abaterei standard* (s) corespunzătoare. Existența corelației între *valoarea medie* și *abaterea standard* este cunoscută sub denumirea de *efect de proporționalitate*. Efectul de *proporționalitate directă* (**Fig.1.19a**) indică faptul că:

- în zonele în care au fost determinate *valori mari* ale caracteristicii studiate *variabilitatea este mare* și ca urmare erorile de estimare vor fi mari;

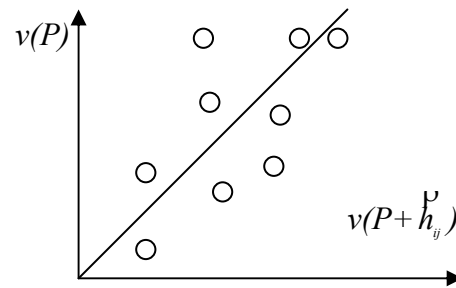


Fig.1.18.Diagrama de continuitate

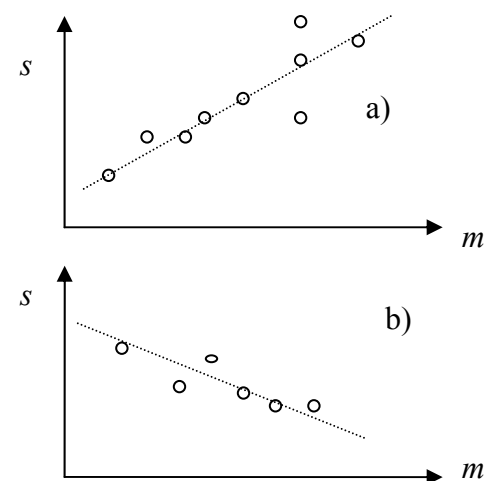


Fig.1.19.Diagrame de variabilitate cu efect de proporționalitate directă(a) și inversă(b).

- în zonele în care au fost determinate *valori mici* ale caracteristicii studiate *variabilitatea este mică* și ca urmare erorile de estimare vor fi mici.

Este frecvent și efectul de *proporționalitate inversă* (**Fig.1.19b**), adică:

- în zonele în care au fost determinate *valori mari* *variabilitatea este mică* și ca urmare erorile de estimare vor fi mici;
- în zonele în care au fost determinate *valori mici* *variabilitatea este mare* și ca urmare erorile de estimare vor fi mari.

Lipsa efectului de proporționalitate nu permite prognozarea mărimii relative a erorilor de estimare. Este situația unei variabilități cu amplitudine mare și cu o distribuție spațială neuniformă care conduce în general la erori mari de estimare a distribuției spațiale a caracteristicii studiate.

Pe parcursul realizării reprezentării grafice a datelor începem să “simțim” caracteristicile structurii spațiale ale variabilei geologice studiate prin intermediul *văzului*. Dintre toate simțurile *văzului* ne face să cunoaștem mai bine și descoperă mai multe diferențe. Dacă îi oferim imagini *reprezentative* pentru structura studiată el ne permite să ne formăm o imagine corectă asupra acesteia.

În etapa reprezentării grafice a datelor ne familiarizăm cu particularitățile unei anumite *structuri spațiale*; altfel spus, acumulăm *experiență* prin confruntare cu *individualul* reducând riscul aplicării greșite a instrumentelor geostatistice în etapele ulterioare de prelucrare.

Analiza în paralel a *hărților punctuale*, *hărților simbolice*, a *diagramelor de continuitate* și a *diagramelor de variabilitate* este absolut necesar să preceadă orice încercare de estimare spațială a unei variabile geologice deoarece elimină riscul interpretărilor greșite determinate de erori grosolane ce pot interveni pe parcursul prelucrărilor:

- *la realizarea fișierelor de date (prin greșeli comise la introducerea valorilor)*;
- *funcționarea greșită a programelor de calcul etc.*

Reprezentările grafice realizate trebuie permanent consultate și confruntate cu rezultatele intermediare și cele finale ale estimărilor distribuției spațiale.

A6.Harta punctuală

Să se construiască harta punctuală necesară analizei variabilității spațiale a porozității unor formațiuni grezoase (în care sunt acumulate hidrocarburi), determinată în 140 de foraje. Cele 140 de valori ale porozității sunt cuprinse între 8 și 40% (tabelul 1.9) și au fost determinate pe un interval de adâncime cuprins între 1800 și 1900m, coordonatele fiind exprimate în unități grafice care nu sunt specificate pentru a păstra “secretul” amplasării zonei cercetate.

Rezolvare:

Sistemul de referință în care este realizată harta punctuală are originea în punctul de coordonate (240, 415) (**Fig.1.20**).

Punctele de observație (cele 140 de foraje) sunt reprezentate prin discuri având diametrul proporțional cu valoarea porozității determinată în fiecare foraj.

O astfel de hartă punctuală (**Fig.1.20**) permite sesizarea următoarelor caracteristici ale rețelei de probare a porozității formațiunilor ce constituie rezervoare de hidrocarburi:

- distribuția punctelor de observație este neuniformă, motiv pentru care gradul de acoperire cu informație a zonei cercetate este mai bun în zona NV-ică (dreptunghiul gri) și foarte slab în NE-ul zonei, unde există puține foraje (dreptunghiul hașurat);
- valorile mari și mici ale porozității se găsesc distribuite uniform în toată zona cercetată, ceea ce indică o mare variabilitate spațială care conduce la erori mari în etapa de estimare a distribuției spațiale a porozității.

Tabelul 1.9. Valorile porozității totale determinate în 140 de foraje

X	Y	n	X	Y	n	X	Y	n	X	Y	n
254	416	32.10	253	426	24.53	248	436	27.50	246	447	35.50
255	418	20.80	249	427	18.50	248	436	24.58	246	447	37.61
259	418	35.70	258	427	29.40	247	436	31.50	246	447	28.75
257	419	17.30	249	427	11.30	248	437	33.43	247	447	34.58
253	419	34.55	250	427	34.80	247	437	24.75	246	447	34.68
263	419	32.53	249	428	10.80	258	437	23.06	246	447	31.30
247	420	24.18	249	428	33.55	248	437	29.40	248	447	30.65
259	420	25.20	253	428	21.10	247	437	27.10	247	447	17.50
259	420	12.90	259	430	31.05	247	437	22.85	255	448	30.00
247	420	25.68	258	430	26.35	248	437	34.00	240	448	27.60
257	421	21.10	258	431	32.30	249	437	32.22	241	448	17.55
247	421	25.42	248	431	33.00	251	437	15.40	246	448	31.55
247	421	32.95	258	432	13.35	249	437	21.93	241	449	32.14
248	422	34.38	258	432	28.82	258	437	33.00	242	449	15.30
260	422	19.65	257	432	25.30	248	438	26.50	246	449	34.00
248	422	31.26	250	434	32.50	258	438	28.12	241	449	20.26
248	422	29.10	248	435	18.35	258	438	24.40	246	449	31.12
262	422	30.17	248	435	22.05	252	438	27.87	246	449	27.75
253	422	33.85	248	435	10.02	249	438	27.20	241	449	18.74
260	423	19.15	248	435	31.47	258	438	22.50	241	449	20.01
254	423	19.00	248	435	27.78	252	441	25.10	241	449	19.89
248	423	16.67	258	435	8.00	249	442	18.00	246	450	38.40
251	423	21.07	248	435	24.25	247	444	33.08	248	450	27.00
262	424	33.90	248	435	29.10	247	445	26.27	246	450	35.55
262	424	33.14	248	435	30.46	248	445	30.50	247	451	33.30
246	424	27.85	248	435	35.10	249	445	32.25	245	451	31.52
251	424	29.67	248	436	29.65	251	445	32.00	245	451	31.10
261	425	24.17	248	436	13.50	246	446	32.70	245	451	26.28
250	425	25.20	248	436	27.10	246	447	34.36	245	452	33.17
251	425	27.78	258	436	27.28	246	447	40.50	245	452	24.90
259	426	32.75	248	436	29.53	246	447	34.68	246	452	32.20
253	426	14.56	248	436	30.13	246	447	37.50	245	452	16.50
253	426	23.00	249	436	27.30	246	447	36.86	246	453	27.05
251	426	22.60	258	436	24.31	246	447	31.80	245	454	19.13
245	454	15.80	245	454	9.86	244	454	39.00	246	454	23.80

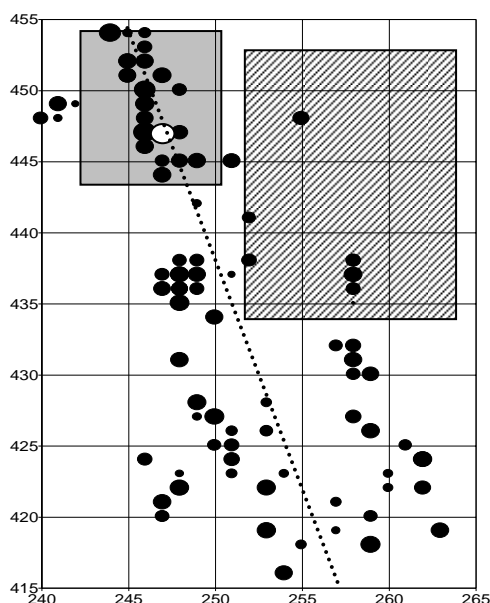


Fig. 1.20. Harta punctuală a porozității

hărții, și valoarea parametrului determinat. Astfel, dacă se realizează o hartă punctuală pentru colțul din dreapta-jos al hărții punctuale din **Fig. 1.20**, aceasta va fi una de detaliu a porozității (**Fig. 1.21**).

În etapa de analiză variografică, hărțile punctuale sunt completate cu raza de influență a modelelor de variogramă, pentru a evalua cantitativ gradul de acoperire cu informație a zonei cercetate.

COMENTARIU

Construirea hărții punctuale a porozității formațiunilor grezoase are de rezolvat două obiective importante:

- să vizualizeze distribuția spațială a punctelor de observație pentru a sesiza gradul de acoperire cu informație a zonei cercetate;
- să permită identificarea zonelor cu porozități favorabile acumulărilor de hidrocarburi (cele cu porozități mari).

Hărțile punctuale servesc mai bine primul obiectiv, cel de-al doilea fiind mai eficient rezolvat prin intermediul hărților simbolice.

A7. Harta simbolică și indicatoare

Să se construiască harta simbolică și cea indicatoare pentru distribuția spațială a valorilor porozității într-o zonă (**Fig. 1.20**) cercetată prin 140 de foraje (**tabelul 1.9**).

Rezolvare:

Elementele pe baza cărora se construiesc cele două tipuri de hărți sunt:

- dimensiunile zonei cercetate:
 - $X_{min} = 240$; $X_{max} = 265$
 - $Y_{min} = 415$; $Y_{max} = 455$

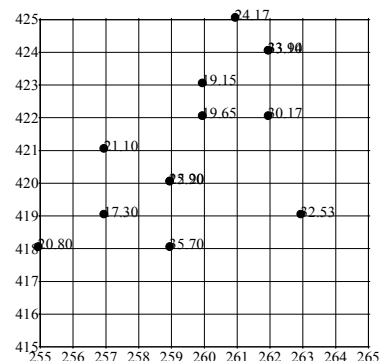


Fig. 1.21. Harta punctuală cu etichete atașate

Dacă numărul punctelor de observație este redus, lângă fiecare dintre acestea se poate scrie, fără să afecteze claritatea

- *numărul de celule* rectangulare în care se împarte zona cercetată, ales în funcție de gradul de detaliu la care vrem să reprezentăm variația porozității în zona respectivă (pentru aplicația realizată, alegerea acestor elemente a avut în vedere obținerea unei hărți simple care să permită identificarea elementelor de calcul):
 - $nx = 11$ (numărul de celule pe un rând)
 - $ny = 17$ (numărul de celule pe o coloană)
- *metoda de calcul* a valorii medii a porozității în fiecare din cele 187 de celule (11x17):
 - s-a utilizat *kriging-ul zonal*
- *clasele de valori* ale porozității asociate simbolurilor
 - pentru *harta simbolică* s-au ales patru clase de valori reprezentate fiecare prin tonuri de gri (**Fig.1.22**):
 - de la 12% la 18%
 - de la 18% la 24%
 - de la 24% la 30%
 - de la 30% la 36%
 - pentru *harta indicatoare* (**Fig.1.23**) s-au ales două clase de valori :
 - de la 12 % la 24%, reprezentată prin gri;
 - de la 24% la 36%, reprezentată prin alb.

Fiecare tip de hartă *simbolică/indicatoare* trebuie să aibă atașată legenda simbolurilor utilizare pentru clasele de valori definite. Pe hărțile realizate (**Fig.1.22** și **Fig.1.23**) sunt figurate și celulele rețelei de discretizare pentru evidențierea gradului de detaliere aplicat.

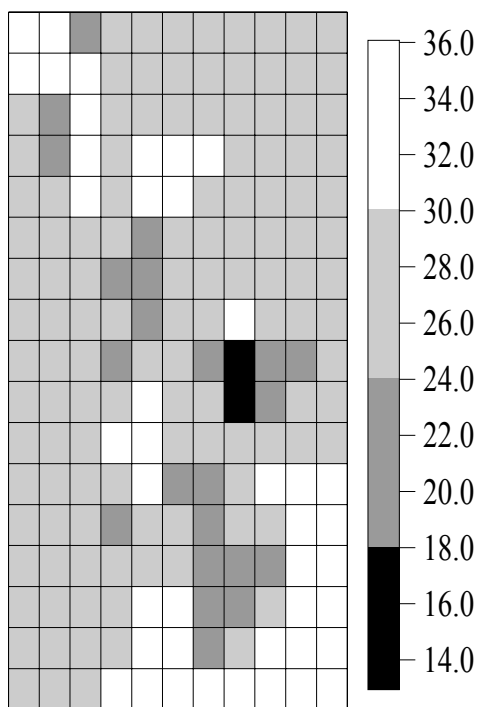


Fig.1.22.Hartă simbolică cu patru simboluri (clase de valori)

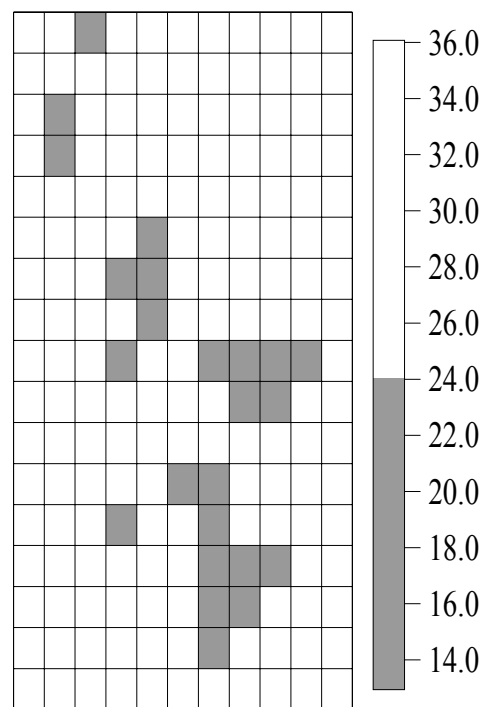


Fig.1.23.Harta indicatoare a porozității (două simboluri)

Harta simbolică permite sesizarea unor caracteristici calitative ale structurii spațiale a porozității:

- zonarea simetrică a porozității de-a lungul direcției NV-SE;
- existența unei zone cu porozități cuprinse între 18 și 24% în partea centrală bordată de o parte și de alta de porozități cuprinse între 24 și 30%;
- apariția unei zone singulare cu porozități reduse (<18%) în partea centrală a zonei studiate.

Harta indicatoare evidențiază :

- direcția NV-SE față de care distribuția porozității este simetrică;
- discontinuitatea pe direcție a zonei cu porozități mai mici de 24%.

COMENTARIU

Atenție!!!...din această etapă de prelucrare intervine subiectivismul cercetătorului.

Cum alegeți numărul de celule rectangulare în care se împarte zona cercetată ?

*În principiu, alegerea se face în funcție de densitatea punctelor de observație: cu cât densitatea punctelor este mai mare cu atât numărul celulelor rectangulare poate fi mai mare (suprafața fiecărei celule mai mică) deoarece avem cu ce calcula valorile medii în fiecare celulă. **Numărul de celule îl putem alege într-un domeniu și nu este o valoare unică.***

Cum alegem numărul claselor de valori asociate simbolurilor ?

*În principiu, în funcție de gradul de detaliere cu care vrem să reprezentăm distribuția variabilei studiate și ținând seama de variabilitatea acesteia reflectată de valorile disponibile: dacă dorim un grad mare de detaliere și variabilitatea caracteristicii studiate este pronunțată se alege un număr mare de clase de valori. **Nu există o formulă care să permită determinarea numărului optim de clase de valori.** Doar în cazul hărților indicatoare pot exista criterii speciale care pot impune valori obligatorii ale celor două clase de valori, cum ar fi conținuturi minime de exploatare pentru anumite minereuri, valori maxime pentru anumiți componenți în apele potabile etc.*

*Am ajuns deja la etapa în care cu aceleași date doi cercetători pot realiza două hărți simbolice diferite fără să putem aprecia care e mai aproape **de adevărul pe care... de cele mai multe ori nu îl vom ști niciodată!!!***

*Cunoașterea distribuției spațiale **adevărate** a caracteristicii studiate necesită **alegerea** unor parametri ce pot minimiza erorile de estimare dar nu le pot elimina. Particularitatea hărții realizate depinde de numărul și tipul parametrilor aleși.*

A8. Diagrama de continuitate

Să se construiască două diagrame de continuitate pentru porozitatea formațiunilor explorate prin 140 de foraje (**tabelul 1.9**), pe direcția $N20^{\circ}V$, pentru distanțele dintre perechile de puncte: $|h_1^p|=0,7$ și $|h_2^p|=22,4$. Pentru calculul numărului de perechi de puncte plasate pe direcția $N20^{\circ}V$, să se utilizeze:

- toleranța de direcție de 10 grade sexagesimale;
- toleranța de distanță de 10% din $|h_i^p|$ ($i=1,2$)

Pentru care dintre cele două distanțe, pe direcția aleasă, continuitatea porozității este mai bună?

Rezolvare:

Din cele 140 de foraje în care au fost determinate porozitățile, au fost identificate:

- 19 perechi situate pe direcția $N20^{\circ}V \pm 10^{\circ}$ și la distanța de 0,7 \pm 0,07;
- 49 perechi situate pe direcția $N20^{\circ}V \pm 10^{\circ}$ și la distanța de 22,4 \pm 2,24.

Tabelul 1.10. Perechile de valori pentru diagramele de continuitate

Direcția $N20^{\circ}V$ (Modul vector = 0,7)			Direcția $N20^{\circ}V$ (Modul vector = 22,4)					
Nr	p(x)	p(x+0,7)	Nr.	p(x)	p(x+22,4)	nr.	p(x)	p(x+22,4)
1	13.50	29.53	1	10.8	15.3	26	23.0	34.7
2	27.50	24.58	2	10.8	17.6	27	23.0	37.5
3	28.12	22.15	3	11.3	17.6	28	23.0	36.9
4	28.75	31.30	4	14.6	17.5	29	23.0	37.6
5	30.13	27.50	5	12.9	25.1	30	23.0	40.5
6	31.80	28.75	6	10.8	27.6	31	24.2	32.0
7	31.80	31.30	7	11.3	27.6	32	25.3	32.2
8	33.08	26.27	8	14.6	28.8	33	26.4	27.0
9	33.43	29.40	9	14.6	31.8	34	25.2	25.1
10	34.36	28.75	10	14.6	31.3	35	27.8	32.7
11	34.00	27.75	11	10.8	32.1	36	27.8	34.4
12	34.36	31.80	12	14.6	34.7	37	27.8	34.7
13	35.50	31.30	13	14.6	35.5	38	27.8	34.6
14	35.50	28.75	14	14.6	37.5	39	27.8	40.5
15	36.86	31.30	15	14.6	36.9	40	32.1	27.2
16	37.61	31.30	16	14.6	37.6	41	33.9	33.1
17	37.50	31.30	17	19.0	33.1	42	34.8	27.6
18	36.86	37.61	18	21.1	38.4	43	35.7	26.5
19	35.10	13.50	19	21.1	27.0	44	35.7	27.2
			20	23.0	17.5	45	33.6	20.3
			21	23.0	28.8	46	33.6	20.0
			22	23.0	31.8	47	33.6	19.8
			23	23.0	31.3	48	33.6	18.7
			24	23.0	34.7	49	34.8	17.6
			25	23.0	35.5			

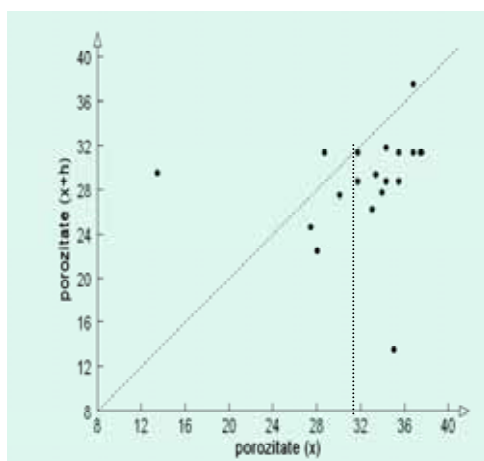


Fig.1.24.Diagrama de continuitate pentru direcția $N20^{\circ}V$ și modulul vectorului de poziție $h = 0,7$.

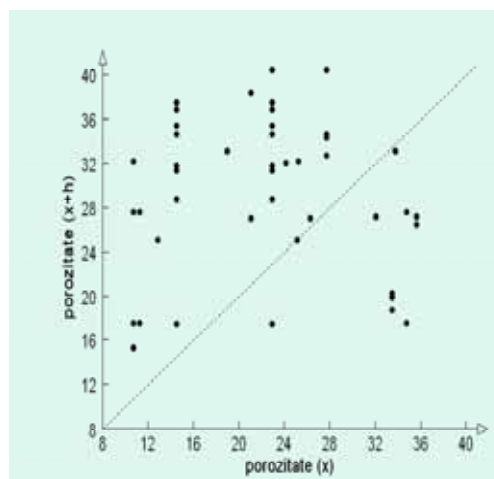


Fig.1.25.Diagrama de continuitate pentru direcția $N20^{\circ}V$ și modulul vectorului de poziție $h = 22,4$.

Valorile porozităților din perechile identificate (**tabelul 1.10**) au fost reprezentate în sistemele de referință rectangulare $[(p(x), p(x+0,7))]$ (**Fig.1.24**) și $[p(x), p(x+22,4)]$ (**Fig.1.25**), reprezentări care constituie *diagramele de continuitate* solicitate în aplicație în care:

- $p(x)$ - porozitatea într-un punct oarecare x ;
- $p(x+0,7)$ - porozitatea într-un punct situat la distanța 0,7 de punctul x , pe direcția $N20^{\circ}V$;
- $p(x+22,4)$ - porozitatea într-un punct situat la distanța 22,4 de punctul x , pe direcția $N20^{\circ}V$.

Aprecierea continuității se face în funcție de intensitatea corelației între perechile de valori concretizată în distribuția geometrică a punctelor. În cazul celor două diagrame construite este evident că gruparea punctelor este mai “strânsă” pentru $|h_1^p| = 0,7$ și deci pentru această distanță *continuitatea este mai bună*.

Altfel spus, estimările spațiale care se vor realiza la distanțe de 0,7 de punctele de observație pe direcția $N20^{\circ}V$ vor fi mai precise decât cele realizate pe aceeași direcție la distanțe de 22,4 de punctele de observație.

Gradul de grupare a punctelor poate fi cuantificat numeric prin intermediul unui coeficient de corelație binară. Dacă se utilizează spre exemplu coeficientul lui Pearson (r), pe baza valorilor din **tabelul 1.10** se obțin valorile:

- $r_1 = 0,15$ pentru direcția $N20^{\circ}V$ și $|h_1^p| = 0,7$;
- $r_2 = -0,08$ pentru direcția $N20^{\circ}V$ și $|h_1^p| = 22,4$,

care confirmă concluzia obținută prin examinarea vizuală a diagramelor de continuitate.

COMENTARIU

“Știința cea mai de sus și mai cu autoritate este aceea care cunoaște scopul în vederea căruia trebuie făcut un lucru”.

Precizez scopul diagramei de continuitate:

Prognoza valorii unei caracteristici geologice pe o anumită poziție în spațiu. Poziția este precizată prin intermediul direcției și distanței pentru care este realizată diagrama de continuitate a acelei caracteristici.

Să aducem rapid lumina clarificatoare a exemplului utilizând:

- *harta punctuală a porozității (Fig.1.20);*
- *diagrama de continuitate a porozității întocmită pentru direcția N20°V și $\left| \frac{p}{h_1} \right| = 0,7$ (Fig.1.24).*

În forajul amplasat în punctul de coordonate $x = 246$, $y = 447$ (discul alb din Fig.1.20), porozitatea determinată este $p = 31,8\%$ (tabelul 1.9). Pe direcția N20°V (marcată în Fig.1.20 printr-o linie punctată) la 0,7 distanță porozitatea poate fi, conform diagramei de continuitate, $p_1 = 28,77\%$ sau $p_2 = 31,30\%$ (tabelul 1.10 și Fig.1.24).

Dacă acceptăm că distribuția spațială a porozității este staționară, la 0,7 distanță măsurată pe direcția N20°V de orice punct în care porozitatea este $p = 31,8\%$ porozitatea are una din valorile $p_1 = 28,77\%$ sau $p_2 = 31,30\%$.

Se poate merge mai departe determinându-se modelul analitic al corelației dintre $p(x)$ și $p(x+0,7)$. Pentru diagrama de continuitate din Fig.1.24, în cazul unui model liniar și la o eroare de genul I $\alpha = 5\%$, se obține ecuația modelului liniar al corelației:

$$p(x+0,7) = 24,32 \pm 14,69 + 0,13 \pm 0,44 \cdot p(x)$$

al cărei coeficient Pearson $r = 0,15$.

Cu acest model se poate prognoza în zona cercetată porozitatea, la distanța 0,7 pe direcția N20°V de orice punct în care porozitatea este cuprinsă între $p_{\min} = 13,5\%$ și $p_{\max} = 35,1\%$, adică domeniul de valabilitate al ecuației stabilite.

A9. Diagrama de variabilitate

Grosimea unui strat de argilă care constituie acoperișul unui acvifer sub presiune din zona unei stații de benzină a fost determinată în 100 de foraje amplasate într-o rețea pătratică cu latura de 100m (tabelul 1.11).

Să se realizeze diagrama de variabilitate a grosimii stratului de argilă și să se estimeze variabilitatea erorilor de estimare a distribuției spațiale a acesteia.

Tabelul 1.11. Grosimea stratului de argilă

Nr	X [m]	Y [m]	Gr. [m]	Nr.	X [m]	Y [m]	Gr. [m]	Nr.	X [m]	Y [m]	Gr. [m]	Nr.	X [m]	Y [m]	Gr. [m]
1	100	100	11.6	26	300	600	6.2	51	600	100	8.3	76	800	600	7.2
2	100	200	10.9	27	300	700	6.1	52	600	200	10.3	77	800	700	5.6
3	100	300	18.2	28	300	800	5.2	53	600	300	12.6	78	800	800	6.1
4	100	400	8.3	29	300	900	4.3	54	600	400	6.9	79	800	900	5.9
5	100	500	6.9	30	300	1000	4.2	55	600	500	8.2	80	800	1000	6.0
6	100	600	6.2	31	400	100	9.9	56	600	600	8.3	81	900	100	6.4
7	100	700	5.4	32	400	200	9.5	57	600	700	7.5	82	900	200	5.3
8	100	800	5.3	33	400	300	10.2	58	600	800	2.2	83	900	300	19.6
9	100	900	5.5	34	400	400	22.3	59	600	900	4.6	84	900	400	10.3
10	100	1000	5.9	35	400	500	10.5	60	600	1000	5.4	85	900	500	7.3
11	200	100	12.3	36	400	600	8.3	61	700	100	8.6	86	900	600	6.2
12	200	200	10.5	37	400	700	6.3	62	700	200	10.6	87	900	700	5.7
13	200	300	19.5	38	400	800	4.2	63	700	300	11.2	88	900	800	5.8
14	200	400	9.5	39	400	900	4.1	64	700	400	8.5	89	900	900	6.2
15	200	500	7.0	40	400	1000	4.2	65	700	500	8.4	90	900	1000	6.1
16	200	600	6.5	41	500	100	12.2	66	700	600	7.3	91	1000	100	19.3
17	200	700	6.2	42	500	200	15.6	67	700	700	6.3	92	1000	200	18.2
18	200	800	5.2	43	500	300	17.3	68	700	800	6.2	93	1000	300	20.3
19	200	900	4.2	44	500	400	12.6	69	700	900	5.0	94	1000	400	12.5
20	200	1000	4.3	45	500	500	9.5	70	700	1000	6.1	95	1000	500	8.3
21	300	100	12.2	46	500	600	12.3	71	800	100	7.5	96	1000	600	5.8
22	300	200	8.3	47	500	700	5.6	72	800	200	9.5	97	1000	700	5.9
23	300	300	15.6	48	500	800	5.3	73	800	300	15.3	98	1000	800	6.2
24	300	400	8.9	49	500	900	5.5	74	800	400	11.2	99	1000	900	4.3
25	300	500	7.3	50	500	1000	6.2	75	800	500	8.3	100	1000	1000	6.3

Rezolvare:

Evaluarea corelației între valoarea medie a grosimii stratului de argilă și abaterea standard specifică fiecărei zone a arealului explorat se bazează pe tehnica ferestrei mobile. Aceasta constă în alegerea unei suprafețe bidimensionale, adică o fereastră mobilă cu care se baleiază zona cercetată. Pentru fiecare poziție a ferestrei mobile pe baza grosimilor stratului de argilă din punctele de observație corespunzătoare se calculează media (m) și abaterea standard (s).

Forma, dimensiunile ferestrei mobile și modalitatea de explorare a zonei cercetate sunt determinate de densitatea punctelor de observație (mare sau mică), distribuția lor spațială (regulată sau neregulată) și gradul de detaliu al analizei:

- dacă distribuția punctelor de observație este regulată, în cazul aplicației în curs de rezolvare se recomandă o fereastră mobilă de formă geometrică regulată:

forma ferestrei mobile alese: pătrată;

- dimensiunile ferestrei mobile trebuie să asigure pe fiecare poziție a acesteia un număr suficient de puncte de observație care să confere reprezentativitate statistică parametrilor calculați:

lățimea ferestrei mobile = 300m

lungimea ferestrei mobile = 300m

- modalitatea de explorare a zonei cercetate asigură gradul de detaliere dorit al variației spațiale pentru medie și abatere standard :

suprapunerea ferestrelor succesive cu o treime din suprafața lor.

Cu parametrii aleși se pot calcula pentru 16 poziții succesive ale ferestrei mobile pătrate cu suprafața de 10000 m² (100x100) valorile mediilor și abaterilor standard ale grosimii. Aceste valori (**tabelul 1.12**) se asociază cu coordonatele centrelor ferestrei mobile pe cele 16 poziții figurate cu un disc cenușiu (**Fig.1.26**). În aceeași figură, punctele de observație sunt reprezentate prin cercuri cu diametrul mai mic și de asemenea sunt figurate două ferestre mobile pe poziții vecine din care se vede zona de suprapunere. Cu cât zona de suprapunere este mai mare cu atât se pierde mai mult din gradul de detaliere al variației spațiale pentru medie și dispersie, realizându-se o “netezire” a variabilității grosimii. La limită, dacă suprafața ferestrei mobile este egală cu suprafața întregii zone cercetate, se obțin o singură medie și o singură dispersie fără a se putea analiza distribuția lor spațială.

Cu ajutorul valorilor din **tabelul 1.12** se pot realiza două reprezentări grafice:

- *harta distribuției spațiale a abaterii standard pentru grosimea stratului de argilă (Fig.1.26);*
- *diagrama de variabilitate a grosimii stratului de argilă (Fig.1.27).*

Tabelul.1.12.Mediile și dispersiile în cele 16 poziții.				
Nr.	X	Y	m	s
1	250	250	12.25	4.21
2	450	250	12.95	4.52
3	650	250	11.25	3.37
4	850	250	12.98	6.22
5	250	450	10.71	5.21
6	450	450	11.06	4.31
7	650	450	10.32	3.06
8	850	450	10.48	4.45
9	250	650	6.43	1.46
10	450	650	7.06	2.48
11	650	650	7.14	2.20
12	850	650	6.66	0.99
13	250	850	5.04	0.82
14	450	850	5.06	1.23
15	650	850	5.59	1.11
16	850	850	5.86	0.53

Harta distribuției abaterii standard (Fig.1.26) indică o variabilitate spațială mai mare în sudul perimetrului explorat, zonă în care valorile abaterii standard sunt mai mari de 3,5 m și o variabilitate spațială mai redusă în nordul acestuia.

Errorile de estimare (ε) pentru distribuția grosimii stratului de argilă, în cazul unei probabilități de 95%, pot fi evaluate cu relația $\varepsilon = \pm 2 \cdot s$ (D.Scrădeanu, 1996). Această estimare este orientativă dată fiind aproximația cu care este evaluată distribuția abaterii standard în această fază.

Diagrama de variabilitate (Fig.1.27) este o diagramă binară în care pe axa absciselor se reprezintă valorile abaterii standard iar pe axa ordonatelor valorile

mediei. Dispoziția celor 16 perechi de valori în acesta diagramă indică un *efect de proporționalitate* directă. Acest lucru arată că în zonele unde valorile grosimii sunt mici abaterea standard este mică și deci și erorile de estimare a distribuției spațiale sunt mici. Analog, în zonele unde sunt valori mari ale grosimii stratului de argilă, abaterea standard este mare și ca urmare erorile de estimare a distribuției acesteia sunt mari.

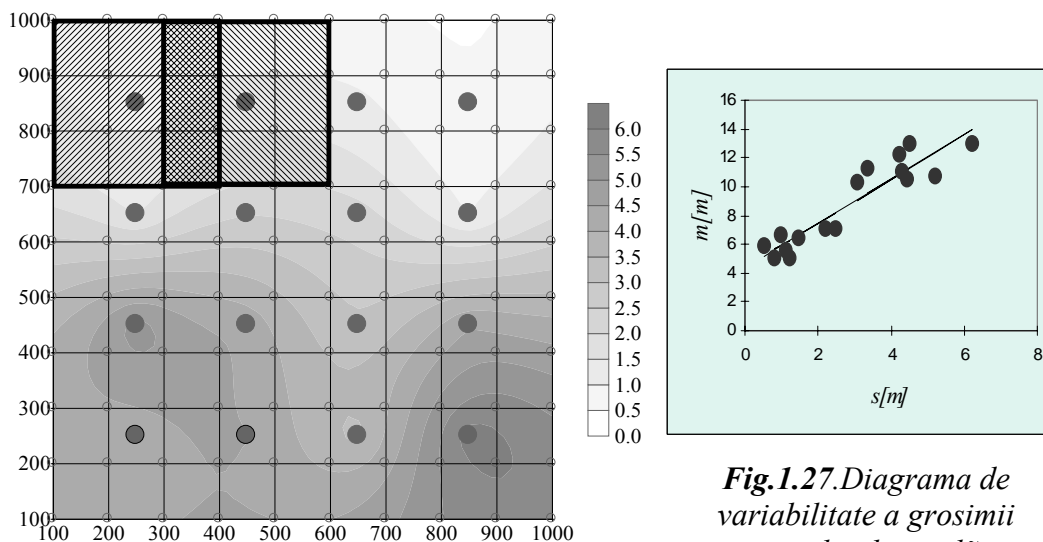


Fig.1.26.Harta distribuției abaterii standard

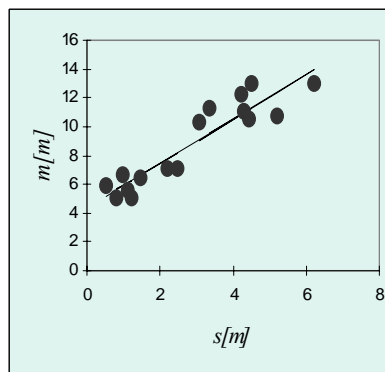


Fig.1.27.Diagrama de variabilitate a grosimii stratului de argilă.

Cele două reprezentări grafice realizate ne permit familiarizarea cu distribuția erorilor de estimare a grosimii în zona cercetată (prin intermediul *hărții distribuției abaterii standard*) și corelarea erorilor de estimare a grosimii cu valoarea estimată (prin intermediul *diagramei de variabilitate*).

COMENTARIU

Diagrama de variabilitate a grosimii stratului de argilă (Fig.1.27) sintetizează o medie a variabilității acesteia în zona cercetată. Ea este expresia variabilității spațiale a caracteristicilor studiate în tot domeniul studiat, filtrată de distribuția spațială a acesteia. Chiar dacă în diagrama de variabilitate nu apare explicit poziția spațială a valorilor caracteristicii studiate ea este implicată de pozițiile succesive ale ferestrei mobile care permit calculul mediilor și dispersiilor “localizate”.

Diagrama de variabilitate contribuie la precizarea relației dintre valoarea estimată și eroarea de estimare.

Cu ajutorul **reprezentărilor grafice** reușim să aflăm “la ce ne putem aștepta” în urma estimărilor spațiale care vor urma în etapele ulterioare, adică:

- ce valori are caracteristica studiată în punctele în care nu a fost probată;
- cât de mari vor fi erorile de estimare ale acestor valori.

Rețineți că informațiile obținute cu ajutorul **reprezentărilor grafice** relativ la structurile spațiale sunt de **natură calitativă**.

1.3.2. ANALIZA PARAMETRICĂ A DATELOR

Pentru analiza parametrică a datelor se schimbă instrumentele de lucru. Renunțăm la pensulă și la culoarea cu care în etapa reprezentării grafice a datelor, prin *limbajul viu*, specific artei, realizăm imaginea *descriptivă* a structurii. Acum se utilizează alte instrumente specifice *limbajului abstract* al științei care permit automatizarea *calculelor* necesare estimării distribuției spațiale. Câștigul este obținut cu prețul pierderii contactului intuitiv cu *realitatea* structurii spațiale. Instrumentele de lucru sunt acum *funcții matematice* care descriu legătura între valorile variabilei studiate și pozițiile lor în spațiu. Mărimile cu care se operează sintetizează caracteristicile morfologice ale structurii în numere și relații matematice care constituie *modelul matematic* al structurii.

Abstractizarea structurii se definitivează în această etapă de *analiză parametrică*. Ce a fost plăcut vederii noastre în faza *reprezentării grafice a datelor* dispare și apar relațiile modelelor matematice, pentru mulți dintre noi inexpressive, dar care dețin magia manipulării numerelor, indispensabilă estimărilor spațiale.

Utilizarea corectă a instrumentelor matematice pe care începem să le construim acum impune cunoașterea și respectarea algoritmilor de prelucrare și a regulilor de interpretare a rezultatelor. La capătul acestor prelucrări ne așteaptă o nouă imagine a structurii estimate (mai detaliată decât cea obținută în hărțile punctuale, hărțile simbolice sau indicatoare) fie sub forma unei *hărți*, fie a unei *secțiuni geologice*. Dacă această imagine, rezultată prin prelucrarea datelor cu ajutorul modelului matematic, este în contradicție flagrantă cu *reprezentările grafice ale datelor* înseamnă că modelul matematic construit nu este adecvat datelor disponibile ori este utilizat incorect.

Să fim foarte atenți la construirea instrumentelor prin care abstractizăm structura. Acestea, pe lângă bagajul matematic cu care sunt înzestrate prin definiție, înglobează și o componentă particulară extrasă din datele a căror prelucrare o realizează (termenul tehnic al procedurii prin care realizează aceste lucru este *calare*).

În mod uzual instrumentele de prelucrare care se utilizează în această etapă de analiză parametrică sunt trei *funcții de distanță*: *funcția de covarianță* ($c(h)$), *corelograma* ($\rho(h)$) și *funcția de variogramă* ($\gamma(h)$).

Toate cele trei funcții de continuitate sunt puternic influențate de *valorile extreme* care nu se încadrează în variabilitatea globală a selecției de date cu care se operează. Dacă forma uneia dintre cele trei funcții nu este clar definită, este de mare utilitate examinarea *diagramelor de continuitate* pentru identificarea acestor valori și eliminarea lor.

Calculate în raport cu o singură variabilă, funcțiile de continuitate permit cuantificarea continuității acesteia în raport cu direcția și distanța.

Funcția de *covarianță* $c(h)$ reprezintă variația similitudinii valorilor din două puncte în raport cu distanța dintre ele. Ea se calculează cu formula:

$$c(h) = \frac{1}{N(h)} \sum_{(i,j) | h_{ij}=h}^{N(h)} v_i v_j - m_{-h} m_{+h} \quad (1.4)$$

în care:

$N(h)$ este numărul perechilor de puncte separate prin vectorul $|\vec{h}|$;

$(i, j)_{|h_{ij}=h}$ - perechea de puncte (p, p_j) separate prin vectorul $|\vec{h}_{ij}|$;

v_i - valoarea variabilei din originea vectorului $|\vec{h}_{ij}|$;

v_j - valoarea variabilei în vârful vectorului $|\vec{h}_{ij}|$;

m_{-h} - media valorilor situate în originea celor $N(h)$ vectori $|\vec{h}|$.

$$m_{-h} = \frac{1}{N(h)} \sum_{i|_{h_{ij}=h}}^{N(h)} v_i \quad (1.5)$$

m_{+h} - media valorilor situate în vârful celor $N(h)$ vectori $|\vec{h}|$.

$$m_{+h} = \frac{1}{N(h)} \sum_{j|_{h_{ij}=h}}^{N(h)} v_j \quad (1.6)$$

Corelograma $(\rho(h))$ este o funcție de covarianță standardizată și se calculează cu relația:

$$\rho(h) = \frac{c(h)}{\sigma_{-h} \sigma_{+h}} \quad (1.7)$$

în care:

σ_{-h} este abaterea standard a tuturor valorilor aflate în originea celor $N(h)$ vectori $|\vec{h}|$:

$$\sigma_{-h}^2 = \frac{1}{N(h)} \sum_{i|_{h_{ij}=h}}^{N(h)} v_i^2 - m_{-h}^2 \quad (1.8)$$

σ_{+h} este abaterea standard a tuturor valorilor aflate în vârful celor $N(h)$ vectori $|\vec{h}|$.

$$\sigma_{+h}^2 = \frac{1}{N(h)} \sum_{j|_{h_{ij}=h}}^{N(h)} v_j^2 - m_{+h}^2 \quad (1.9)$$

Funcția de variogramă $(\gamma(h))$ reprezintă variația varianței erorii de estimare în raport cu distanța dintre punctul în care se cunoaște valoarea variabilei și cel în care aceasta se estimează. Altfel spus, valoarea variogramei pentru un anumit vector $|\vec{h}|$ exprimă eroarea care se comite atunci când se atribuie variabilei în punctul $p+h$ valoarea sa din punctul p . Ea se calculează cu relația:

$$\gamma(h) = \frac{1}{2 \cdot N(h)} \sum_{(i,j)|_{h_{ij}=h}}^{N(h)} (v_i - v_j)^2 \quad (1.10)$$

Toate cele trei funcții univariate de distanță nu sunt afectate de sensul vectorului $|\vec{h}|$, fiind funcții pare; dacă se schimbă indicele i cu j în toate formulele de calcul ale funcțiilor de distanță valorile $c(h)$, $\rho(h)$ și $\gamma(h)$ nu se schimbă:

$$c(h) = c(-h); \quad \rho(h) = \rho(-h); \quad \gamma(h) = \gamma(-h) \quad (1.11)$$

Ideea de continuitate poate fi extinsă și la două variabile. Corelația spațială între două variabile u și v depinde de continuitatea fiecăreia. Continuitatea poate fi exprimată grafic printr-o *diagramă de continuitate* în care pe cele două axe se reprezintă $v(p)$ și $u(p + |\vec{h}|)$. Este evident că pentru $|\vec{h}| = 0$ într-o diagramă de continuitate bivariată (u, v) nu toate punctele se află pe bisectoarea unghiului făcut de cele două axe de coordonate, așa cum se întâmplă în cazul diagramelor de continuitate univariate.

Aceleași funcții, utilizate pentru o singură variabilă, se utilizează cu modificările corespunzătoare pentru descrierea continuității bivariate. Pentru evidențierea aspectului bivariat le vom numi: *funcția de intercovarianță*, *funcția de intercorelație* și *funcția de intervareogramă* și vom introduce doi indici corespunzători celor două variabile $(c_{uv}(h), \rho_{uv}(h), \gamma_{uv}(h))$.

Funcția de intercovarianță se calculează cu relația:

$$c_{uv}(h) = \frac{1}{N(h)} \sum_{(i,j) \mid h_{ij}=h}^{N(h)} u_i v_j - m_{u-h} m_{v+h} \quad (1.12)$$

în care:

u_i - sunt valorile variabilei u ;

v_j - valorile variabilei v ;

$N(h)$ - numărul perechilor de puncte separate prin vectorul h ;

m_{u-h} - media valorilor variabilei u situate în originea celor $N(h)$ vectori.

$$m_{u-h} = \frac{1}{N(h)} \sum_{i \mid h_{ij}=h}^{N(h)} u_i \quad (1.13)$$

m_{v+h} - media valorilor variabilei v situate în vârful celor $N(h)$ vectori:

$$m_{v+h} = \frac{1}{N(h)} \sum_{j \mid h_{ij}=h}^{N(h)} v_j \quad (1.14)$$

Funcția de intercorelație este dată de ecuația:

$$\rho_{uv}(h) = \frac{c_{uv}(h)}{\sigma_{u-h} \sigma_{v+h}} \quad (1.15)$$

în care:

σ_{u-h} - abaterea standard a tuturor valorilor variabilei u , aflate în originea celor $N(h)$ vectori h :

$$\sigma_{u-h}^2 = \frac{1}{N(h)} \sum_{i \mid h_{ij}=h}^{N(h)} u_i^2 - m_{u-h}^2 \quad (1.16)$$

σ_{v+h} - abaterea standard a tuturor valorilor aflate în vârful celor $N(h)$ vectori $|\vec{h}|$:

$$\sigma_{v+h}^2 = \frac{1}{N(h)} \sum_{j \mid h_{ij}=h}^{N(h)} v_j^2 - m_{v+h}^2 \quad (1.17)$$

Funcția de intervariogramă ($\gamma_{uv}(h)$) se calculează cu relația:

$$\gamma_{uv}(h) = \frac{1}{2N(h)} \sum_{(i,j) \mid h_{ij}=h}^{N(h)} (u_i - u_j) \cdot (v_i - v_j) \quad (1.18)$$

Funcțiile de continuitate bivariate nu sunt în totalitate funcții pare. Funcțiile de intercovarianță și de intercorelație depind de sensul vectorului $|\vec{h}|$ pe o anumită direcție ($c_{uv}(h) \neq c_{uv}(-h)$; $\rho_{uv}(h) \neq \rho_{uv}(-h)$), în timp ce funcția de intervariogramă nu depinde de sensul vectorului h ($\gamma_{uv}(h) = \gamma_{uv}(-h)$).

A10. Covarianța, corelograma și variograma

Să se calculeze covarianța, corelograma și variograma pentru grosimile stratului de argilă care protejează acviferul sub presiune din zona unei stații de benzină (tabelul 1.11), pentru direcția NS și distanțele 100, 200, 300, 400, 500, 600 și 700 m.

Rezolvare:

Calculul funcțiilor de distanță pe direcția NS în cazul unei rețele regulate de puncte de observație (Fig.1.28) simplifică mult procedeul de calcul. La un număr redus de puncte de observație calculul poate fi realizat chiar manual, fără ajutorul unui program de calcul, experiment care permite acomodarea cu aceste instrumente matematice.

Pentru explicitarea calculului valorii covarianței pe direcția NS, corespunzător distanței de 100 m conform formulei (1.4), în Fig.1.28 sunt reprezentate câteva poziții ale vectorului de poziție \vec{h} orientat NS și de modul 100 m.

În colțul din stânga jos sunt reprezentate primele trei poziții succesive corespunzătoare primilor trei termeni ai sumei (primele trei perechi de valori situate pe direcția NS la distanța de 100 m); în stânga sus - ultima poziție de pe prima coloană pentru al nouălea termen al sumei, iar sus în dreapta - ultima poziție pentru ultimul termen al sumei care pentru aceste poziții are forma:

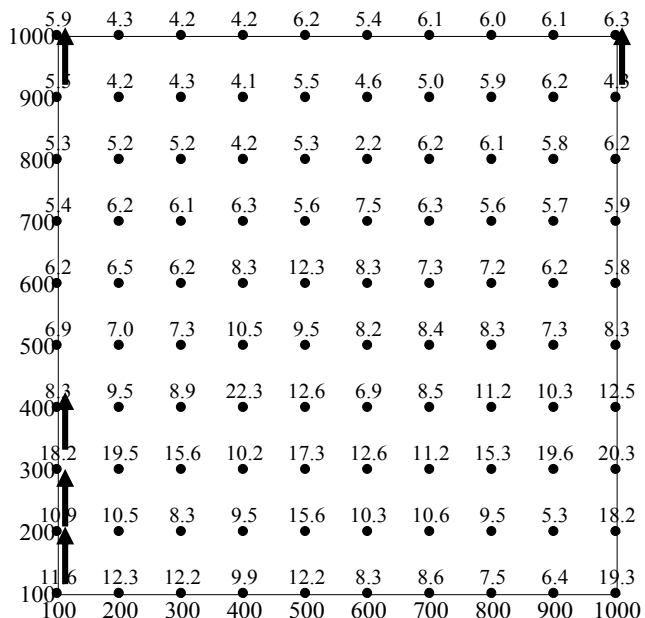


Fig.1.28. Rețeaua de puncte de observație pentru calculul covarianței, corelogramei și variogramei.

$$\sum_{(i,j) \in \mathcal{I}_{h_{ij}=h}}^{N(h)} v_i v_j = 11,6 \cdot 10,9 + 10,9 \cdot 18,2 + 18,2 \cdot 8,3 + \dots + 5,5 \cdot 5,9 + \dots + 4,3 \cdot 6,3 = 7438,88 \quad m^2$$

Baleind tot arealul cercetat cu vectorul \vec{h} se identifică $N(h) = 90$ de perechi de valori situate pe direcția NS la distanța de 100 m între ele.

Media celor 90 de valori situate în originea celor 90 de poziții ale vectorului \vec{h} este:

$$m_{-h} = \frac{1}{N(h)} \sum_{i \in \mathcal{I}_{h_{ij}=h}}^{N(h)} v_i = \frac{1}{90} (11,6 + 10,9 + 18,2 + \dots + 5,5 + \dots + 4,3) = 8,37 \quad m$$

iar media celor 90 de valori situate în vârful celor 90 de poziții ale vectorului \vec{h} este:

$$m_{+h} = \frac{1}{N(h)} \sum_{j \in \mathcal{I}_{h_{ij}=h}}^{N(h)} v_j = \frac{1}{90} (10,9 + 18,2 + 8,3 + \dots + 5,9 + \dots + 6,3) = 8,62 \quad m$$

Valoarea *covarianței* pentru distanța de 100 m pe direcția NS este deci:

$$c(h) = \frac{1}{N(h)} \sum_{(i,j) \in \mathcal{I}_{h_{ij}=h}}^{N(h)} v_i v_j - m_{-h} m_{+h} = \frac{1}{90} \cdot 7438,88 - 8,37 \cdot 8,62 = 10,47 \quad m^2$$

Calculul *corelogramei* pentru aceeași distanță implică calculul dispersiilor celor 90 de valori din originea și vârful celor 90 de poziții ale vectorului \vec{h} :

$$\sigma_{-h}^2 = \frac{1}{N(h)} \sum_{i \in \mathcal{I}_{h_{ij}=h}}^{N(h)} v_i^2 - m_{-h}^2 = \frac{1}{90} (11,6^2 + 10,9^2 + 18,2^2 + \dots + 5,5^2 + \dots + 4,3^2) - 8,37^2 = 14,78 \quad m^2$$

$$\sigma_{+h}^2 = \frac{1}{N(h)} \sum_{j \in \mathcal{I}_{h_{ij}=h}}^{N(h)} v_j^2 - m_{+h}^2 = \frac{1}{90} (10,9^2 + 18,2^2 + 8,3^2 + \dots + 5,9^2 + \dots + 6,3^2) - 8,62^2 = 17,62 \quad m^2$$

Corelograma pentru direcția NS și distanța de 100 m este:

$$\rho(h) = \frac{c(h)}{\sigma_{-h} \sigma_{+h}} = \frac{10,47}{\sqrt{14,78} \cdot \sqrt{17,62}} = 0,65$$

Calculul *variogramei* pe direcția NS și pentru același modul de 100 m nu implică decât valorile măsurate pentru cele 90 de poziții ale vectorului \vec{h} :

$$\gamma(h) = \frac{1}{2 \cdot N(h)} \sum_{(i,j) \in \mathcal{I}_{h_{ij}=h}}^{N(h)} (v_i - v_j)^2 = \frac{1}{2 \cdot 90} [(11,6 - 10,9)^2 + \dots + (4,3 - 6,3)^2] = 5,76 \quad m^2$$

În mod analog, pentru celelalte distanțe se obțin valorile (*tabelul 1.13*) pe baza cărora se pot construi reprezentările grafice ale funcțiilor de *covarianță* (*Fig.1.29*) și *variogramă* (*Fig.1.30*). Graficul *corelogramei* are aceeași morfologie cu al *funcției de covarianță* dar cu o amplitudine unitară, deoarece *corelograma* este o *covarianță normată*.

Tabelul 1.13. Valorile funcțiilor de distanță

Nr.	N(h)	h	$\gamma(h)$	c(h)	$\rho(h)$	m _{-h}	M _{+h}	σ^2_{-h}	σ^2_{+h}
1	90	100	5.76	10.47	0.65	8.37	8.62	14.78	17.62
2	80	200	7.47	8.45	0.53	8.43	8.64	14.47	17.33
3	70	300	7.96	8.75	0.53	8.45	8.75	15.34	17.99
4	60	400	7.36	9.59	0.57	8.56	8.72	17.20	16.66
5	50	500	5.95	11.47	0.66	8.79	8.42	18.80	15.91
6	40	600	7.09	10.98	0.61	8.43	8.67	18.43	17.66
7	30	700	5.00	14.14	0.76	8.26	8.95	15.80	22.00

Toate cele trei funcții de distanță sintetizează caracteristicile structurale ale variabilei studiate într-o formă care poate fi modelată matematic și care permite automatizarea calculelor necesare estimării spațiale.

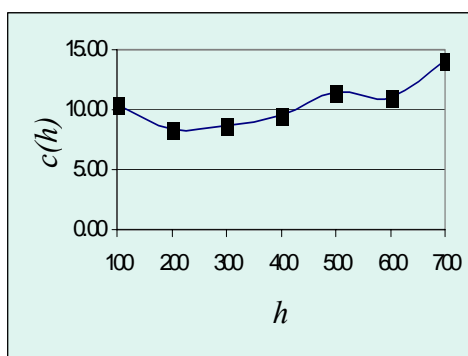


Fig.1.29.Funcția de covarianță

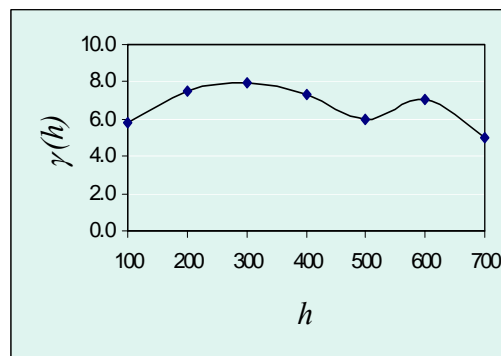


Fig.1.30.Funcția de variogramă

Disponerea “ordonată” a valorilor funcțiilor, cu evidențierea unei corelații funcționale între valoarea funcției și distanță, sugerează întotdeauna existența unei legități de variație spațială a caracteristicii studiate, legitate care este baza estimării spațiale.

COMENTARIU

Existența unei corelații între valoarea funcției de distanță, oricare ar fi ea (covarianță, corelogramă sau variogramă), și distanță este dovada structurării spațiale a variabilei studiate. Altfel spus, pe baza datelor disponibile se poate identifica o lege de variație spațială a variabilei studiate. Legea de variație spațială se utilizează pentru calculul structurilor spațiale și al erorilor asociate.

Lipsa unei astfel de corelații funcționale, evidențiată prin variația aleatoare în raport cu distanța a valorilor funcțiilor, indică lipsa unei legități de variație spațială și în consecință lipsa instrumentului care să asigure controlul erorilor în etapa de

estimare a distribuției spațiale a caracteristicii studiate. În astfel de situații, pentru zona cercetată se apelează la modelele statistice cu care se evaluează **o valoare unică** a caracteristicii studiate, ca **medie statistică** a tuturor valorilor disponibile. Orice încercare de a evalua distribuția spațială a caracteristicii studiate în aceste condiții (finalizată de exemplu printr-o hartă cu izolinii) este un exercițiu fără semnificație deoarece erorile cu care sunt construite aceste hărți pot depăși cu un ordin de mărime valorile estimate.

Utilizarea unui program de tipul SURFER-ului pentru realizarea unei hărți cu izolinii trebuie precedată de identificarea legii de variație spațială a caracteristicii ce urmează a fi cartografiată.

Fără cunoașterea legii de variație spațială harta cu izolinii construită pentru o variabilă este afectată de erori a căror valoare nu poate fi cunoscută, deci ea este inutilizabilă pentru estimări cantitative !

Dacă ai valoarea unei variabile în **doar trei puncte** și cel mai performant program de tip SURFER, nu-l face de râs folosindu-l pentru construirea hărții conturale a acelei variabile.

A11. Intercovarianța, intercorelograma și intervareograma

Care este corelația spațială între porozitatea unui acvifer sub presiune și a grosimii stratului de argilă care îl protejează împotriva poluării produse de o stație de benzină ?

Determinarea celor doi parametri - porozitatea formațiunilor acvifere și grosimea stratului de argilă din acoperiș - este realizată într-o rețea pătratică de 100 de puncte de observație (**tabelul 1.14**).

Rezolvare:

Corelația spațială între cei doi parametri poate contribui în mod considerabil la creșterea preciziei estimării distribuției spațiale a acestora precum și la reducerea costului măsurătorilor executate în teren.

Creșterea preciziei se realizează prin creșterea numărului de condiționări în etapa de estimare spațială. Reducerea costului este posibilă prin reducerea numărului de puncte de observație în care se determină ambele variabile cercetate, una dintre ele putând fi determinată pe baza valorilor celeilalte utilizându-se funcțiile de intercorelație identificate.

Corelația spațială între porozitate (P2 în **tabelul 1.14**) și grosime (Gr. în **tabelul 1.14**) se poate cerceta prin intermediul funcțiilor de intercovarianță, intercorelogramă și intervareogramă.

Deoarece existența unei corelații spațiale este condiționată de corelația globală între cele două variabile, înaintea calculului funcțiilor de intercorelație este eficientă o analiză, cel puțin calitativă, a acestei corelații prin intermediul diagramei de corelație.

Pentru a argumenta cele afirmate mai sus se calculează în continuare funcțiile de intercovarianță, intercorelogramă și intervareogramă pe direcția NS pentru distanțele de 100, 200, 300, 400, 500, 600 și 700 m în două variante:

- Varianta I: pentru grosimea ecranului protector și porozitatea acviferului determinate prin măsurători (Gr și P2);
- Varianta a II-a: pentru grosimea ecranului determinată prin măsurători și o porozitate "generată" matematic în corelație liniară cu grosimea (Gr și P1).

Varianta I

Suportul grafic pentru exemplificarea calculelor este *harta punctuală* în care fiecărui punct de observație îi sunt atașate valorile grosimii ecranului protector și porozitățile acviferului determinate în cele 100 de puncte ale unei rețele pătratice cu parametrul 100 m (**Fig.1.31**). Deasupra punctelor de observație sunt scrise valorile grosimilor ecranului protector (*Gr.*) iar sub fiecare punct sunt scrise valorile porozității acviferului (*P2*).

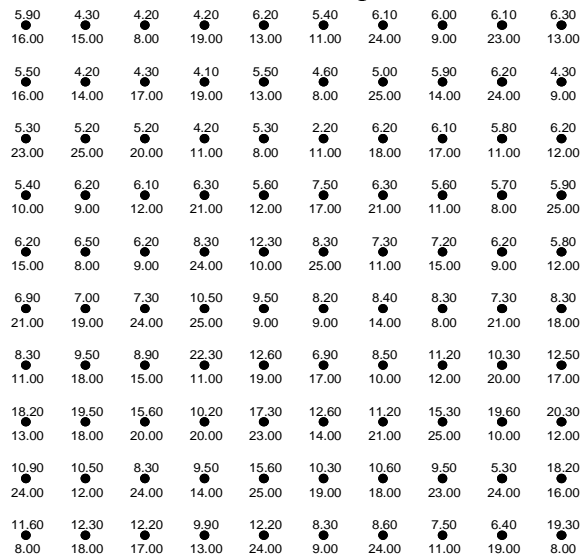


Fig.1.31. Harta punctuală a grosimii ecranului protector și a porozității acviferului.

În prima etapă se construiește *diagrama de corelație* între grosimea ecranului protector (*Gr*) și a porozității determinate prin măsurători (*P2*) (**Fig.1.32**). Diagrama de corelație evidențiază independența dintre cele două variabile luate în considerare. Acest rezultat nu recomandă trecerea la calculul funcțiilor de intercorelație deoarece lipsa unei corelații globale induce și lipsa unei corelații spațiale.

Pentru argumentarea celor afirmate se calculează în *continuare funcția de intervareogramă* pe direcția NS și pentru distanțele 100, 200, ..., 700 m. Explicarea calculului pentru prima valoare a funcției de variogramă pentru direcția NS și distanța 100 m este:

$$\gamma_{Gr,P2}(100) = \frac{1}{2N(h)} \sum_{(i,j)_{h_{ij}=h}}^{N(h)} (Gr_i - Gr_j) \cdot (P2_i - P2_j) =$$

$$= \frac{1}{2 \cdot 90} [(11,6 - 10,5)(8,6 - 24,0) + \dots + (4,3 - 6,3)(9,0 - 13,0)] = -1,05,$$

în care sunt scriși doar primul și ultimul termen al sumei ce conține 90 de termeni, corespunzând celor 90 de perechi de puncte de observație situate pe direcția NS la distanța de 100 m.

Rezultatele calculelor pentru celelalte valori ale *funcției de intervareogramă* precum și ale celor de *intercovarianță* și *intercorelație* sunt sintetizate în **tabelul 1.15**.

Reprezentarea grafică a valorilor funcției de *intervareogramă* (**Fig.1.33**) prin dispoziția haotică a punctelor indică lipsa unei corelații spațiale între grosimea ecranului protector (*Gr*) și porozitatea acviferului (*P2*). În concluzie, estimarea distribuției spațiale a grosimii ecranului protector nu poate fi ameliorată pe baza cunoașterii distribuției spațiale a porozității acviferului.

Varianta a II-a

Pentru a evidenția efectul corelației globale asupra intercorelației spațiale s-a generat o serie de 100 de valori de porozitate (PI) aflate în *corelație liniară perfectă* cu grosimea ecranului protector (Gr) (**Fig.1.34**). Pe baza hărții punctuale din **Fig.1.36** și a acelorași relații de calcul se obțin valorile din **tabelul 1.16** precum și graficul *intervariogramei* (**Fig.1.35**) care evidențiază clar existența unei corelații spațiale între grosimea ecranului protector și porozitatea PI .

În această variantă, evaluarea distribuției spațiale a grosimii ecranului protector poate fi realizată cu *aceeași precizie* pe baza unui *număr mai mic de valori* ale grosimii la care să se adauge: valori ale porozității determinate în acele puncte în care nu se cunoaște grosimea, precum și *funcția de intervareogramă* care leagă cele două variabile (Gr , PI).

Metoda de estimare spațială care integrează acest tip de informații poartă numele de *cokriging*. Ea permite evaluarea distribuției spațiale pentru variabile a căror determinare este foarte costisitoare (ex.: pesticide) prin intermediul altora cu care sunt corelate spațial și au un preț de cost redus al determinărilor (ex.: granulozitate, porozitate etc.).

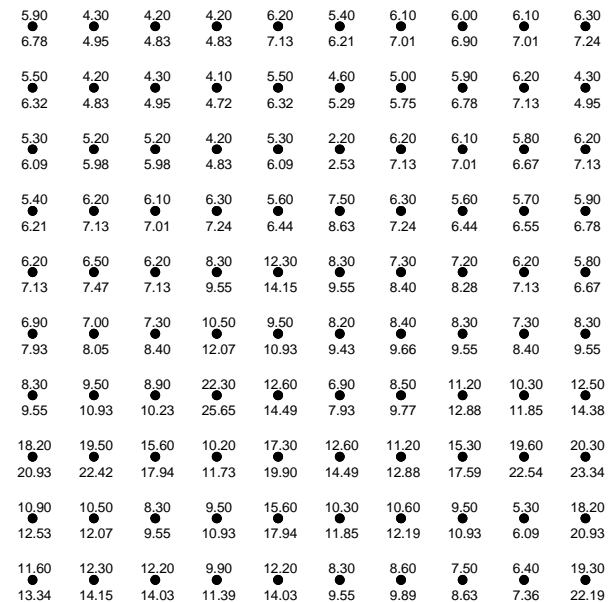


Fig.1.36. Harta punctuală a grosimii ecranului protector și a porozității generate matematic (PI).

COMENTARIU

Analiza parametrică a datelor ne aruncă în lumea abstractă a instrumentelor matematice. Există două variante pentru a depăși cu succes acest real șoc: o bună cunoaștere a regulilor după care operează aceste instrumente sau însușirea modului de utilizare a unor programe de calcul prietenoase care utilizează aceste instrumente.

*Diviziunea muncii în cercetarea științifică fundamentală sau aplicată este dusă la extrem. Realizarea unui program de calcul complex, acum în pragul mileniului trei, este rezultatul colaborării unei echipe complexe formată din: **matematicieni**, care rezolvă schemele de calcul ale programului, **programatori** care programează schemele de calcul ale matematicienilor, **graficieni** care asigură designul interfeței grafice cu utilizatorii programului, **psihologi** care armonizează relațiile programului cu utilizatorii (le fac mai prietenoase). Programele de calcul utilizate în analiza parametrică a datelor elimină dificultățile utilizării instrumentelor matematice.*

Pentru utilizarea corectă a acestor programe este necesar să se știe:

- *modul de organizare în fișiere a datelor primare;*
- *obiectivul prelucrării (identificarea unei legi experimentale a structurii spațiale);*
- *modul de evitare a obstacolelor din calea atingerii obiectivului prelucrării;*
- *limitele instrumentelor utilizate.*

*La acest prim pas pe calea abstractizării structurilor spațiale obstacolele din calea identificării legii structurilor spațiale rezultă din **analiza variabilității globale** și a **reprezentărilor grafice**.*

Trebuie să vi le reamintesc? Desigur!!!

***Primul obstacol** este “anormalitatea” distribuției frecvenței valorilor prelucrate.*

Dacă normalitate nu e, nimic nu e !

Și ați reținut cu siguranță că dacă dispunem de un program de normalizare a distribuției frecvențelor unui set de valori îl utilizăm chiar dacă este posibil ca acele valori să fie distribuite normal. Paza bună...!

***Al doilea obstacol** este prezența valorilor extreme !*

Față de extremisme, nici o îndurare !

Întâi le identificăm, utilizând diagramele de probabilitate și de continuitate după care le eliminăm progresiv pe parcursul calculului funcțiilor de distanță până la ameliorarea corelației între distanțe și valorile funcției.

Gradul de ameliorare a corelației include o notă de subiectivism care distinge prelucrările unui începător de cele ale maestrului. Dacă se greșește în această fază și nu se alege varianta optimă, în etapele ulterioare vom primi semnale care ne obligă să revenim și să facem corecturile necesare. Un începător poate munci o săptămână fără să poată trece cu succes de această etapă pe când un maestru alege varianta optimă fără ezitare.

Singura rețetă pentru a ajunge ... maestru în analiza parametrică a datelor este prelucrarea unui volum mare de date cu variabilități complexe.

1.3.3. ANALIZA STAȚIONARITĂȚII

Descrierea, chiar de manieră exhaustivă a structurilor (spațiale sau temporare), obținută pe baza unei densități mari a punctelor de observație nu poate rezolva problemele estimării spațiale, deoarece “valorile numerice nu sunt *realitatea*, ci doar o primă imagine, analitic foarte bogată, dar structural foarte săracă a acesteia” (Matheron, G., 1978).

Descrierea spațială nu asigură înțelegerea fenomenului de o manieră care să permită explicarea genezei datelor disponibile sau prognoza evoluției acestora între punctele de observație. Pentru realizarea acestor obiective este necesară depășirea stadiului unei interminabile și sterile prelucrări a datelor brute și realizarea pasului următor de construire a unor *modele* care să adauge informație suplimentară. Recurgerea la *modelele topo-probabiliste* presupune două niveluri de abstractizare progresivă: *variabila regionalizată și funcția aleatoare*.

Variabila regionalizată este o funcție de punct ($f(p)$) cu ajutorul căreia se realizează descrierea fenomenelor regionalizate. Prin statutul de funcție matematică, variabila regionalizată câștigă vis-à-vis de fenomenul pe care îl descrie o anumită autonomie, ea putând *să-și trăiască propria sa viață*. Pentru obiectivele unui studiu practic este inutil de subliniat că această libertate a *variabilei regionalizate* trebuie strict supravegheată în interesul interpretării corecte a rezultatelor prelucrărilor.

La acest prim nivel de abstractizare, realizat prin *variabila regionalizată*, este poate momentul semnalării problematicei revenirii de la model la realitatea fizică a modelului. Corectitudinea prelucrărilor matematice întreprinse nu garantează sensul fizic al rezultatelor. La fiecare etapă este necesar să se facă distincția între proprietățile modelului matematic și cele ale realității fizice pe care acesta o reprezintă.

Din punct de vedere *matematic* o variabilă regionalizată este o funcție $f(p)$, p fiind un punct de coordonate (x, y, z) într-un spațiu tridimensional finit. Această precizare este deosebit de importantă, ea subliniind faptul că *modelele topo-probabiliste (geostatistice)* nu au caracter de universalitate, ele fiind valabile într-un anumit domeniu spațial, cel definit de punctele în care se determină valorile variabilei studiate.

Din punct de vedere geostatistic există posibilitatea unei aproximări directe a variabilei regionalizate, instrumentele utilizate fiind elementare și subordonate noțiunii de *integrală spațială*. În practică, variabila regionalizată fiind disponibilă printr-un eșantionaj finit, aceste integrale se reduc la *sume finite*.

Lucrul direct asupra variabilelor regionalizate prezintă avantajul absenței ipotezelor de natură probabilistă de tipul staționarității și ergodicității. *Geostatistica tranzitivă* este acea parte a Geostatisticii care operează direct asupra variabilei regionalizate.

Variabilitatea spațială a fenomenelor regionalizate, foarte complexă, interzice de multe ori, din punct de vedere practic, studiul matematic direct al variabilei regionalizate. O reprezentare corectă a variabilității fenomenelor regionalizate trebuie să ia în considerare două aspecte aparent contradictorii ale variabilei regionalizate:

- *aspectul general* structurat, care face apel la reprezentări funcționale;
- *aspectul local* aleator care face apel la noțiunea de variabilă aleatoare.

Instrumentul care permite luarea în considerare atât a aspectului structurat, cât și a celui aleator, este *funcția aleatoare*. Ea realizează al doilea nivel de abstractizare și, într-un anumit mod, ne mai îndepărtează de realitatea fizică.

La acest nivel de abstractizare variabila regionalizată ($f(p)$) este considerată o realizare a unei funcții aleatoare. Reprezentarea probabilistă a variabilei regionalizate nu este unică și ea se justifică numai în măsura în care permite o caracterizare simplă a structurii și un formalism omogen și operațional de soluționare a evaluărilor spațiale.

Funcția aleatoare este un ansamblu de *variabile aleatoare* cu poziții distincte într-un spațiu oarecare și a căror dependență este specificată printr-un mecanism probabilist.

O variabilă aleatoare este o variabilă care ia un anumit număr de valori, conform unei anumite legi de repartiție. Cota nivelului hidrostatic dintr-un acvifer freatic $v(p_1)=125$ m, măsurată într-un puț, poate fi considerată ca o realizare particulară a unei anumite variabile aleatoare $V(p_1)$, implantată în punctul p_1 .

Ansamblul cotelor nivelului hidrostatic, pentru toate puțurile implantate într-un acvifer freatic, poate fi considerat ca o realizare particulară a ansamblului de variabile aleatoare: $\{V(p), \text{punctele } p \text{ aparținând acviferului}\}$.

Această definiție a funcției aleatoare înglobează ambele aspecte ale variabilei regionalizate:

- *aspectul local*: în punctul p_1 , $V(p_1)$ este o variabilă aleatoare;
- *aspectul regional*: $V(p)$ este o funcție aleatoare, în sensul că pentru toate cuplurile de puncte p_1 și p_{1+h} variabilele aleatoare corespunzătoare sunt corelate, aceste corelații exprimând structura spațială a variabilei regionalizate.

Interpretarea probabilistă a unei variabile regionalizate ca o realizare particulară a unei funcții aleatoare nu are sens operațional dacă nu putem identifica *legea de probabilitate* ce caracterizează întreg ansamblul de realizări ale funcției aleatoare.

Dificultatea este datorată faptului că nu este posibilă deducerea riguroasă a legii unei funcții aleatoare plecând de la o realizare unică $v(p)$, limitată și ea la un număr finit de implantări ale punctelor p_i . Deoarece de cele mai multe ori se dispune de o realizare unică, limitată la un număr finit de puncte, pentru a face operațional modelul funcției aleatoare este necesară acceptarea unei anumite omogenități statistice spațiale asigurată de *staționaritatea* și *ergodicitatea* fenomenului regionalizat studiat.

În practică, pe un domeniu spațial limitat, fenomenele regionalizate pot fi frecvent considerate omogene, variabila regionalizată $v(p)$ repetându-se în spațiu după aceleași legități. Repetabilitatea în spațiu echivalează de o anumită manieră cu un număr mai mare de realizări ale funcției aleatoare, acest lucru permițând inferența statistică.

Două valori experimentale $v(p_i)$ și $v(p_{i+h})$, implantate în două puncte diferite, pot fi considerate ca două realizări diferite ale aceleiași funcții aleatoare $V(p_i)$. Acest compromis permite inferența *legii de distribuție a funcției aleatoare* $V(p)$ plecând de la *histograma* datelor ($v(p_i)$) și a *speranței matematice* $E\{V(p)\}$, de la *media aritmetică* a valorilor selecției de date disponibile ($m(p)$).

La nivelul modelelor topo-probabiliste *staționaritatea* unui fenomen regionalizat este definită ca *invarianța legii spațiale la translație*. Altfel spus:

Legea spațială, relevantă de un ansamblu de puncte, este staționară dacă nu depinde de poziția acestor puncte.

Modelele topo-probabiliste, mai precis geostatistica lineară, se bazează pe primele două momente ale legii de distribuție pentru definirea diferitelor ipoteze de staționaritate. Pentru geostatistica lineară, două funcții aleatoare $V_1(p)$ și $V_2(p)$ care admit aceleași momente de ordinul unu și doi nu se diferențiază una de alta și sunt considerate ca unul și același model.

Momentele legilor spațiale ale funcțiilor aleatoare utilizate sunt:

$E\{V(p)\}$ - speranța matematică

$$E\{V(p)\} = m(p) \quad (1.19)$$

$Var\{V(p)\}$ - varianța

$$Var\{V(p)\} = E\{[V(p) - m(p)]^2\} \quad (1.20)$$

$c(p_1, p_2)$ - covarianța

$$c(p_1, p_2) = E\{[V(p_1) - m(p_1)][V(p_2) - m(p_2)]\} \quad (1.21)$$

$2\gamma(p_1, p_2)$ - variograma

$$2\gamma(p_1, p_2) = Var\{V(p_1) - V(p_2)\} \quad (1.22)$$

Covarianța și variograma sunt funcții dependente de două implantații p_1 și p_2 iar calculul lor necesită mai multe realizări ale cuplului $\{V(p_1), V(p_2)\}$. Cum acest lucru nu este posibil, deoarece de cele mai multe ori dispunem de o singură serie de măsurători în fiecare punct de probare, numai dacă aceste funcții ar depinde doar de vectorul h (care separă două puncte p_1 și p_2) inferența lor ar fi posibilă. În această ipoteză toate cuplurile $\{V(p_k), V'(p_k)\}$ plasate la distanța $|h|$ pe direcția vectorului h pot fi considerate ca realizări diferite ale cuplului $(V(p_1), V(p_2))$.

Pentru distribuția spațială a conținutului unui poluant, a gradului de nebulozitate a atmosferei, a compoziției granulometrice a unui depozit deltaic, corelația între valorile măsurate în două puncte situate la o anumită distanță constituie o caracteristică intrinsecă a structurilor respective. Această intuiție fizică se traduce în contextul modelelor topo-probabiliste prin ipoteze de *staționaritate la diferite niveluri de rigoare*.

Staționaritatea strictă corespunde unei omogenități statistice foarte avansate. Este greu de găsit un fenomen care să se conformeze aceleiași legi structurale în toate punctele domeniului său spațial.

Staționaritatea de ordinul doi a unei funcții aleatoare este asigurată de:

- existența speranței matematice și independența acesteia de punctul de implantare p :

$$E\{V(p)\} = m, \quad \forall p \quad (1.23)$$

- existența covarianței și invarianța acesteia la translație:

$$c(h) = E\{V(p) \cdot V(p+h)\} - m^2, \quad \forall p \quad (1.24)$$

Existența și staționaritatea covarianței implică existența staționarității varianței și variogramei. Se deduc imediat relațiile:

$$\text{Var}\{V(p)\} = E\{[V(p) - m]^2\} = c(0), \forall p \quad (1.25)$$

și

$$\gamma(h) = \frac{1}{2} E\{[V(p+h) - v(p)]^2\} = c(0) - c(h), \forall p \quad (1.26)$$

În ipoteza de staționaritate de ordinul doi, covarianța și variograma sunt echivalente pentru caracterizarea autocorelației între două variabile $V(p+h)$ și $V(p)$ aflate la extremitățile vectorului h . În aceste condiții se definește și corelograma:

$$\rho(h) = \frac{c(h)}{c(0)} = 1 - \frac{\gamma(h)}{c(0)} \quad (1.27)$$

Existența funcției de variogramă reprezintă o ipoteză mai puțin dură decât existența funcției de covarianță. Există numeroase fenomene regionalizate care nu au nici covarianță, nici varianță finită, dar au variogramă finită. În consecință, se poate lărgi cadrul staționarității de ordinul doi doar la existența variogramei.

Staționaritatea intrinsecă impune cele mai lejere condiții:

- existența speranței matematice și independența ei de punctul de implantare:

$$E\{V(p)\} = m, \forall p \quad (1.28)$$

- independența autocovarianței creșterilor față de translație și dependența ei de h :

$$\text{Var}\{V(p+h) - V(p)\} = E\{[V(p+h) - V(p)]^2\} = 2\gamma(h) \quad (1.29)$$

Trebuie remarcat că ipoteza intrinsecă a staționarității nu antrenează staționaritatea de ordinul doi.

Modelarea topo-probabilistă a variabilelor regionalizate implică luarea în considerare a unui factor total absent din formalismele probabiliste: scara structurii. Variabila regionalizată poate fi considerată sau nu ca o realizare a unui proces staționar în funcție de scara de lucru.

Cu excepția fenomenelor autoomotetice același obiect poate fi considerat regulat sau neregulat, structurat sau nestructurat, staționar sau nestaționar, în funcție de scara la care este studiat.

În practică, variograma și covarianța sunt utilizate pentru distanțe limitate: $h < r$ în care se păstrează omogenitatea statistică. Două variabile $V(p_k)$ și $V(p_{k+h})$ aflate la distanțe $h > r$ nu pot fi considerate ca aparținând aceluiași tip de structură, ele nefiind două realizări ale unui proces staționar, în particular, speranțele lor matematice fiind diferite:

$$E\{V(p_k)\} \neq E\{V(p_k + h)\}, |h| > r \quad (1.30)$$

Pentru astfel de situații se utilizează *funcțiile structurale* $c(p, p+h)$ sau $\gamma(p, p+h)$, care sunt staționare local, pe distanțe mai mici decât r . Această limitare la distanțe $|h| < r$ a ipotezei de staționaritate de ordinul doi (sau staționaritate intrinsecă, dacă se presupune numai existența variogramei) corespunde ipotezei de cvasistaționaritate (sau staționaritate cvasiintrinsecă).

În mod concret o funcție aleatoare este cvasistaționară de ordinul 2 dacă:

- speranța matematică $E\{V(p)\}$ există și este o funcție regulată și cu variație lentă în raport cu poziția punctului la scara rețelei de probare disponibilă;
- covarianța există și este o funcție dependentă de vectorul h_{ij} , și de poziția celor două puncte p_i și p_j :

$$c(p_1 - p_2, p_1, p_2) = E\{[V(p_1) - m(p_1)][V(p_2) - m(p_2)]\}, \quad i=1, \quad j=2 \quad (1.31)$$

La scara informației disponibile, adică pentru poziții p_i și p_j nu foarte îndepărtate, covarianța poate fi considerată funcție de un singur argument și anume distanța dintre cele două puncte $|h_{ij}|$.

Din punct de vedere practic se pot defini vecinătăți mobile în interiorul cărora speranța matematică și covarianța pot fi considerate staționare. Ipoteza de cvasistaționaritate este rezultatul compromisului dintre dimensiunea r a omogenității statistice a fenomenului și densitatea informației disponibile deoarece, pentru atingerea staționarității, reducerea dimensiunii r este limitată doar de volumul de date minim necesar realizării inferenței.

Adoptarea modelului topo-probabilist presupune acordarea setului de date disponibil atributul de *reprezentativ*, ceea ce, din punct de vedere probabilist, echivalează cu proprietatea de *ergodicitate*.

Prin definiție, un proces *staționar* este *ergodic* (satisface ipoteza de ergodicitate) dacă seria mediei spațiale:

$$V_n = \frac{1}{S_n} \int_{S_n} V(p) dp \quad (1.32)$$

converge la speranța matematică $E\{V(p)\}$ (care este invariantă în spațiu, conform ipotezei de staționaritate) atunci când domeniul S_n tinde la infinit.

În condițiile ergodicității este posibil ca, plecând de la observarea variației în spațiu a unui fenomen regionalizat, pe baza unei *realizări unice*, să se deducă legea de distribuție spațială a ansamblului tuturor realizărilor posibile, dar necunoscute. Altfel spus, *ergodicitatea* face să coincidă mediile calculate pe ansamblul realizărilor funcției aleatoare cu mediile spațiale, obținute din valorile măsurate în rețeaua punctelor de observație.

În practica, deoarece se dispune de cele mai multe ori de o singură serie de măsurători în rețeaua punctelor de observație, ergodicitatea nu poate fi testată, ea fiind acceptată ca premiză teoretică a utilizării modelului funcției aleatoare.

Un proces regionalizat, care verifică ipotezele de *staționaritate* și *ergodicitate*, este un proces *omogen*, a cărui modelare poate beneficia de funcția aleatoare și de tot arsenalul de facilități ale acestui instrument probabilist.

Ce instrumente avem la dispoziție pentru a testa staționaritatea structurii spațiale a unei variabile pe baza unui număr finit de valori ale acesteia, determinate într-o rețea de puncte de observație ?

Cum putem aprecia dacă valorile disponibile pot fi prelucrate cu instrumentele matematice ale geostatisticii pentru a obține o estimare spațială corectă ?

Rar găsești prin cărți preocupări pentru tratarea explicită a acestui aspect, el fiind de cele mai multe ori subînțeles, ca și acela al *normalității* distribuției. Lipsa unei astfel de preocupări a determinat ignorarea frecventă a analizei staționarității care determină în etapa de evaluare introducerea *erorilor sistematice*.

Consacrăm tratării staționarității două aplicații în care utilizăm instrumentele *grafice* și *analitice* disponibile pentru realizarea unei imagini intuitive a staționarității.

Tratarea grafică și analitică a staționarității presupune: *identificarea*, *calculul* și *separarea tendințelor* de variație spațială la scară mare și se poate realiza cu ajutorul *suprafețelor polinomiale de tendință*. Rezultatul acestor prelucrări este reducerea erorilor sistematice din etapa de evaluare a distribuției spațiale.

Identificarea tendințelor distribuției spațiale se realizează cu suprafețe polinomiale de tendință:

- *plane* de forma:

$$v_{tend} = A + B \cdot x + C \cdot y \quad (1.33)$$

- *curbe* definite prin *polinoame* de diferite grade, cel mai frecvent fiind utilizate cele de gradul al doilea:

$$v_{tend} = A + B \cdot x + C \cdot y + D \cdot x^2 + E \cdot x \cdot y + F \cdot y^2 \quad (1.34)$$

Suprafața topografică, suprafața piezometrică a unui acvifer sau orice suprafață structurală (hărți conturale cu izocore, izobate, izopahite etc.) au trei dimensiuni:

- *două dimensiuni* geografice reprezentate prin *coordonatele x și y* ale punctelor de observație;
- *a treia dimensiune* este *valoarea variabilei* analizate (cota terenului pentru suprafața topografică, cota nivelului piezometric pentru suprafața piezometrică etc.).

Calculul tendințelor unei variabile se face prin ajustarea suprafețelor de tendință la variația acesteia, variație evidențiată prin valorile măsurate într-o rețea de puncte de observație. Densitatea punctelor de observație și repartiția lor spațială sunt determinante pentru identificarea corectă a suprafețelor de tendință. Este ușor de intuit că o *densitate mare* și o *distribuție uniformă* a punctelor de observație permite identificarea unor *tendințe reprezentative* pentru variabila analizată.

Criteriul de ajustare este *minimizarea* sumei pătratelor abaterilor *valorilor măsurate* (v_{mas_i}) față de cele *calculate* cu ajutorul ecuației suprafeței de tendință (v_{tend_i}). Pentru o suprafață de gradul întâi, de ecuație: $v_{tend} = A + B \cdot x + C \cdot y$, abaterea într-un punct de observație “i” este:

$$\Delta v_i = v_{mas_i} - v_{tend_i} = v_{mas_i} - (A + B \cdot x_i + C \cdot y_i) \quad (1.35)$$

Suma pătratelor abaterilor valorilor calculate pe baza ecuației suprafeței de tendință (v_{tend_i}) de la cele măsurate (v_{mas_i}), pentru toate punctele de observație (n), scrisă sub forma unei funcții în raport cu variabilele A , B și C are forma:

$$F(A, B, C) = \sum_{i=1}^{i=n} (v_{mas_i} - v_{tend_i})^2 = \sum_{i=1}^{i=n} [v_{mas_i} - (A + B \cdot x_i + C \cdot y_i)]^2 \quad (1.36)$$

Valorile coeficienților A , B , și C , conform *criteriului celor mai mici pătrate*, se calculează astfel încât să se obțină valoarea *minimă* a funcției $F(A, B, C)$. Pentru minimizarea funcției $F(A, B, C)$ în raport cu variabilele A , B și C este necesar ca:

$$\frac{\partial F}{\partial A} = \frac{\partial F}{\partial B} = \frac{\partial F}{\partial C} = 0 \quad (1.37)$$

adică:

$$\begin{cases} \frac{\partial F}{\partial A} = \sum_{i=1}^{i=n} 2 \cdot [v_{mas_i} - (A + B \cdot x_i + C \cdot y_i)](-1) = 0 \\ \frac{\partial F}{\partial B} = \sum_{i=1}^{i=n} 2 \cdot [v_{mas_i} - (A + B \cdot x_i + C \cdot y_i)](-x_i) = 0 \\ \frac{\partial F}{\partial C} = \sum_{i=1}^{i=n} 2 \cdot [v_{mas_i} - (A + B \cdot x_i + C \cdot y_i)](-y_i) = 0 \end{cases} \quad (1.38)$$

După efectuarea calculului se obțin cele trei ecuații normale a căror soluție sunt valorile celor trei coeficienți A , B și C :

$$\begin{cases} A \cdot n + B \cdot \sum_{i=1}^{i=n} x_i + C \cdot \sum_{i=1}^{i=n} y_i = \sum_{i=1}^{i=n} v_{mas_i} \\ A \cdot \sum_{i=1}^{i=n} x_i + B \cdot \sum_{i=1}^{i=n} x_i^2 + C \cdot \sum_{i=1}^{i=n} x_i \cdot y_i = \sum_{i=1}^{i=n} v_{mas_i} \cdot x_i \\ A \cdot \sum_{i=1}^{i=n} y_i + B \cdot \sum_{i=1}^{i=n} x_i \cdot y_i + C \cdot \sum_{i=1}^{i=n} y_i^2 = \sum_{i=1}^{i=n} v_{mas_i} \cdot y_i \end{cases} \quad (1.39)$$

adică:

$$\begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^{i=n} x_i & \sum_{i=1}^{i=n} y_i \\ \sum_{i=1}^{i=n} x_i & \sum_{i=1}^{i=n} x_i^2 & \sum_{i=1}^{i=n} x_i \cdot y_i \\ \sum_{i=1}^{i=n} y_i & \sum_{i=1}^{i=n} x_i \cdot y_i & \sum_{i=1}^{i=n} y_i^2 \end{bmatrix}^{-1} \times \begin{bmatrix} \sum_{i=1}^{i=n} v_{mas_i} \\ \sum_{i=1}^{i=n} v_{mas_i} \cdot x_i \\ \sum_{i=1}^{i=n} v_{mas_i} \cdot y_i \end{bmatrix} \quad (1.40)$$

Separarea tendințelor are ca obiective evaluarea *puterii de ajustare* a funcției de tendință și *semnificația statistică* a componentelor acesteia.

O măsură a *puterii de ajustare* (P_ε) a suprafeței de tendință este reducerea procentuală a abaterilor pusă pe seama suprafeței ajustate și calculată în raport cu media valorilor măsurate (\bar{v}_{mas}):

$$P_\varepsilon = 100 \left[1 - \frac{\sum_{i=1}^{i=n} (v_{mas_i} - v_{tend_i})^2}{\sum_{i=1}^{i=n} (v_{mas_i} - \bar{v}_{mas})^2} \right] \quad (1.41)$$

O ajustare perfectă a unei funcții de tendință la datele de observație are $P_\varepsilon=100\%$ și ea presupune lipsa abaterilor valorilor măsurate de la funcția de tendință. Dacă puterea ajustării unei funcții de tendință este mică ($P_\varepsilon=15-20\%$), aceasta este un indiciu că variația spațială a variabilei nu este reprezentată de funcția de tendință aleasă.

Semnificația statistică a funcției de tendință se evaluează prin intermediul analizei dispersionale bifactoriale (D., Scărădeanu, 1995). Instrumentul utilizat pentru testarea semnificației statistice este factorul Fischer experimental care pentru o funcție de tendință cu K termeni și n puncte de observație se calculează cu relația:

$$F_{exp} = \frac{\sum_{i=1}^{i=n} (v_{tend_i} - \bar{v}_{mas_i})^2}{K} \cdot \frac{n - K - 1}{\sum_{i=1}^{i=n} (v_{mas_i} - v_{tend_i})^2} \quad (1.42)$$

Semnificația statistică a funcției de tendință *plană*, pentru care a fost exemplificat modul de calcul al componentelor datorate regresiei liniare ($(B \cdot x + C \cdot y)$; $K=2$), este apreciată pe baza relației dintre factorul Fischer experimental (F_{exp}) și cel teoretic ($F(K, n-K-1, \alpha)$; α - riscul erorii de genul I) calculat pe baza funcției Fischer și disponibil în formă tabelară. Dacă la un risc al erorii de genul I ales $F_{exp} > F(K, n-K-1, \alpha)$, rezultă că suprafața de tendință determinată are semnificație statistică, altfel spus:

Puterea de ajustare evaluată (P_ε) este credibilă în condițiile variabilității spațiale reflectate de cele n valori disponibile.

Tentația eliminării tendințelor pe baza unor polinoame de grade superioare (cinci, șase) este mare datorită puterilor de ajustare mari care se obțin.

Temperarea unei astfel de atitudini este absolut necesară deoarece obiectivul final este estimarea valorilor caracteristicilor studiate *în alte puncte decât în cele probate*. Ori este știut că între punctele de observație suprafețele de tendință conduc la variații proporționale cu gradul polinomului utilizat și nu cu variabilitatea reală a caracteristicii studiate.

A12. Analiza staționarității cu ajutorul variogramei

Să se studieze caracterul distribuției spațiale a conductivității hidraulice a unui acvifer sub presiune constituit din nisipuri siltice, argiloase și fine explorate prin 130 de foraje.

Distribuția forajelor de explorare este neuniformă (Fig.1.37) iar valorile conductivității sunt cuprinse între $4,6 \times 10^{-6}$ și $1,2 \times 10^{-2}$ cm/s (tabelul 1.17).

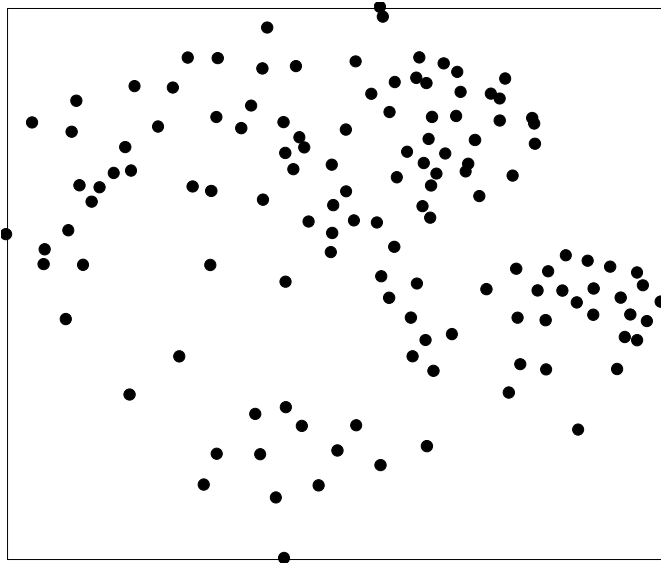


Fig.1.37 Distribuția forajelor de explorare a acviferului sub presiune.

Rezolvare:

Pentru vizualizarea caracterului staționar sau nestaționar al distribuției spațiale a conductivității hidraulice se utilizează în mod curent două metode:

- calculul variogramelor pe principalele direcții structurale;
- reprezentarea variației conductivității hidraulice de-a lungul unor trasee rectilinii.

Instrumentul complet pentru investigarea staționarității este variograma de suprafață care permite estimarea variabilității spațiale în funcție de direcție și distanță. Variograma de suprafață se obține prin reprezentarea variogramelor direcționale (calculate pe mai multe direcții) sub forma unei hărți cu izolinii (Fig.1.38).

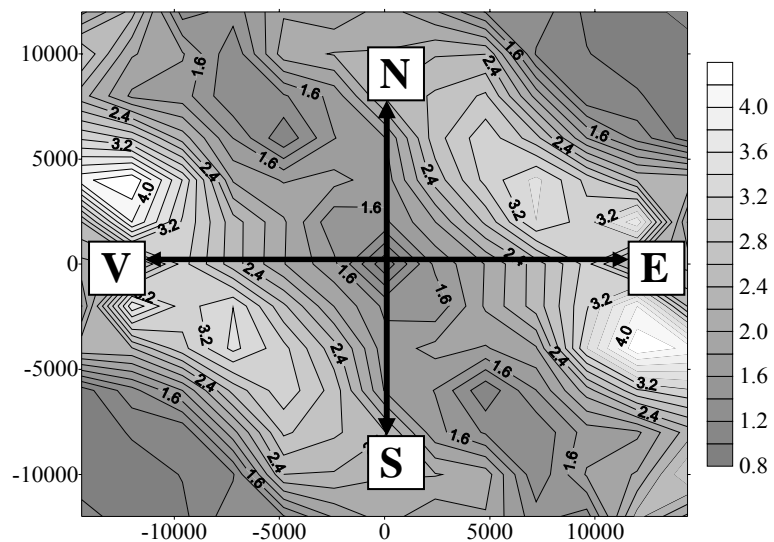


Fig.1.38 Variograma de suprafață a conductivității hidraulice

Variograma de suprafață permite vizualizarea modificării valorilor funcției de variogramă în raport cu direcția și distanța perechilor de puncte.

Tabel 1.17. Conductivitățile hidraulice ale acviferului sub presiune.

Nr.	x[km]	y[km]	k[cm/s]	Nr.	x[km]	y[km]	k[cm/s]	Nr.	x[km]	y[km]	k[cm/s]
1	45.00	80.15	4.6E-06	45	57.67	77.59	3.6E-04	89	45.65	62.79	1.3E-03
2	41.98	75.46	1.1E-05	46	62.59	70.39	3.7E-04	90	54.71	64.36	1.3E-03
3	50.87	73.02	1.4E-05	47	47.94	64.02	3.7E-04	91	63.25	68.66	1.3E-03
4	45.95	74.72	1.6E-05	48	53.18	70.38	3.9E-04	92	54.85	73.64	1.4E-03
5	48.05	74.37	1.6E-05	49	53.49	75.28	4.0E-04	93	61.49	70.76	1.4E-03
6	51.74	73.53	1.8E-05	50	45.19	74.90	4.1E-04	94	60.80	70.19	1.4E-03
7	50.91	74.15	3.0E-05	51	48.22	81.37	4.3E-04	95	53.19	77.94	1.6E-03
8	54.13	68.00	3.5E-05	52	47.73	65.66	4.4E-04	96	58.40	69.58	1.6E-03
9	54.69	79.11	3.7E-05	53	58.99	77.69	4.5E-04	97	41.41	74.88	1.6E-03
10	49.90	73.48	4.4E-05	54	44.39	78.93	4.9E-04	98	59.10	76.65	1.7E-03
11	51.43	74.71	4.4E-05	55	49.39	79.80	5.0E-04	99	60.35	72.11	1.8E-03
12	58.34	71.56	4.5E-05	56	59.21	70.68	5.2E-04	100	50.85	75.79	1.8E-03
13	48.96	76.27	4.6E-05	57	39.18	72.36	5.5E-04	101	39.14	71.75	1.9E-03
14	43.78	77.34	5.1E-05	58	48.90	59.80	5.7E-04	102	52.83	63.58	2.1E-03
15	49.72	76.49	5.6E-05	59	57.89	79.30	5.7E-04	103	62.16	71.65	2.2E-03
16	40.59	74.96	6.1E-05	60	55.10	75.43	5.9E-04	104	55.90	77.78	2.3E-03
17	56.84	74.52	6.9E-05	61	54.53	74.10	6.0E-04	105	59.64	71.46	2.3E-03
18	49.28	75.61	7.2E-05	62	46.17	64.04	6.0E-04	106	51.42	77.22	2.4E-03
19	53.39	72.46	7.9E-05	63	50.31	62.76	6.1E-04	107	51.84	65.20	2.4E-03
20	54.06	69.58	1.1E-04	64	50.80	72.24	6.1E-04	108	44.65	68.00	2.5E-03
21	51.82	79.99	1.1E-04	65	59.55	67.47	6.4E-04	109	55.46	76.25	2.5E-03
22	47.57	78.19	1.2E-04	66	48.03	79.70	6.5E-04	110	54.78	76.84	2.6E-03
23	62.76	68.79	1.2E-04	67	54.31	70.96	6.6E-04	111	41.09	74.29	2.6E-03
24	57.67	78.48	1.3E-04	68	48.57	62.27	7.1E-04	112	56.08	78.75	2.6E-03
25	48.88	77.53	1.4E-04	69	63.65	69.43	7.2E-04	113	57.31	78.68	2.7E-03
26	47.17	77.28	1.4E-04	70	48.97	71.04	7.5E-04	114	52.45	78.68	2.9E-03
27	46.21	80.12	1.5E-04	71	54.66	68.66	8.0E-04	115	56.39	75.83	2.9E-03
28	38.67	77.51	1.6E-04	72	58.51	67.68	8.0E-04	116	59.54	69.48	3.0E-03
29	52.86	71.26	1.6E-04	73	40.28	77.14	8.0E-04	117	60.86	65.03	3.3E-03
30	55.73	68.90	1.6E-04	74	40.74	71.72	8.1E-04	118	62.98	69.70	3.3E-03
31	42.69	75.55	1.7E-04	75	52.80	82.21	8.1E-04	119	56.67	76.80	3.8E-03
32	49.53	76.91	2.0E-04	76	54.58	75.86	8.4E-04	120	42.45	76.51	3.9E-03
33	45.91	71.72	2.1E-04	77	53.90	76.32	8.7E-04	121	52.68	73.44	4.3E-03
34	54.98	67.41	2.3E-04	78	61.47	69.69	9.3E-04	122	63.49	70.89	4.3E-03
35	51.08	64.17	2.3E-04	79	62.44	67.49	9.4E-04	123	54.92	77.73	4.6E-03
36	49.63	65.17	2.3E-04	80	42.83	78.99	9.8E-04	124	63.25	71.41	4.6E-03
37	46.16	77.73	2.6E-04	81	37.61	72.97	1.0E-03	125	55.94	79.57	5.3E-03
38	56.29	75.52	2.9E-04	82	52.92	81.81	1.0E-03	126	57.13	70.73	5.6E-03
39	59.07	77.47	2.9E-04	83	53.40	79.15	1.1E-03	127	40.03	69.52	6.1E-03
40	40.14	73.13	2.9E-04	84	54.40	80.15	1.1E-03	128	54.28	79.33	7.1E-03
41	58.04	66.53	3.0E-04	85	61.25	71.89	1.2E-03	129	58.19	75.35	1.1E-02
42	48.98	65.94	3.0E-04	86	55.40	79.91	1.2E-03	130	64.22	70.23	1.2E-02
43	42.63	66.45	3.1E-04	87	54.88	74.95	1.2E-03				
44	60.22	70.68	3.4E-04	88	40.47	78.39	1.3E-03				

Interpretarea variogramei de suprafață pentru stabilirea caracterului distribuției spațiale se bazează pe analiza variației valorilor funcției de variogramă:

- *stabilizarea* valorilor funcției de variogramă pe o anumită direcție, peste o anumită distanță între perechile de puncte, evidențiază *caracterul staționar* al structurii spațiale;
- *creșterea continuă* a valorilor funcției de variogramă proporțional cu distanța pe o anumită direcție semnalează *caracterul nestaționar* al structurii.

Analiza variogramei de suprafață a conductivității hidraulice (**Fig.1.38**) evidențiază două direcții principale de variabilitate:

- direcția N45°E, de-a lungul căreia se remarcă o creștere proporțională cu distanța (h) a valorilor variogramei, deci un caracter *nestaționar* al structurii;
- direcția N135°V, de-a lungul căreia valorile variogramei cresc până la valori ale distanței de maximum 4500 m după care se stabilizează în jurul valorii de 2 (**Fig.1.39**, **tabelul 1.18**).

Tabel 1.18.

Dist. h[m]	Variograma	
	N135V	N45V
1000	1.41	1.00
1500	0.76	1.50
2000	1.60	1.69
2500	1.33	2.03
3000	1.36	1.08
3500	1.53	2.56
4000	1.90	1.65
4500	1.50	3.00
5000	1.90	2.10
5500	1.68	2.44
6000	1.98	2.99
6500	1.28	2.42
7000	1.37	3.40
7500	1.24	3.08
8000	1.16	3.18
8500	1.28	3.54

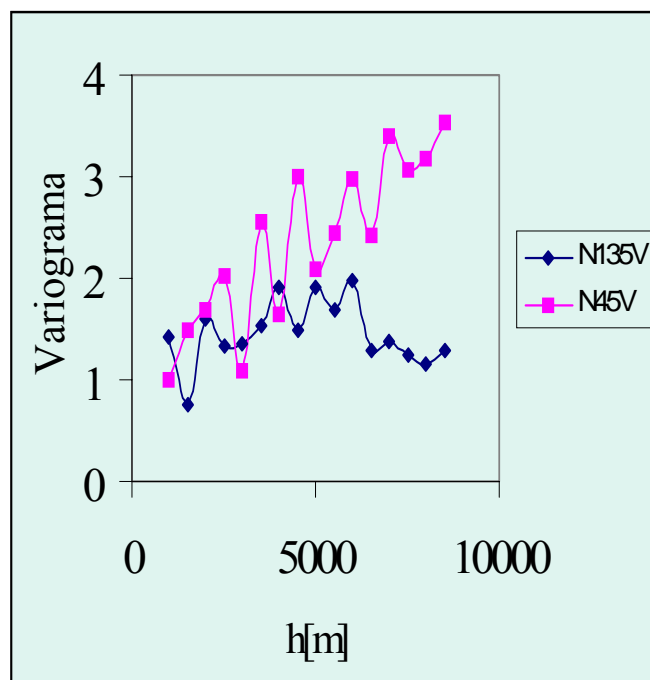


Fig.1.39. Variogramele direcționale pentru conductivitatea hidraulică

Manifestarea diferită a variabilității spațiale în funcție de distanță adaugă structurii și caracterul *anizotrop*. Tratarea anizotropiei este o problemă de mare complexitate și importanță pentru evaluarea corectă a distribuției spațiale și îi vom acorda o atenție specială în cadrul *analizei variografice*.

Pentru estimarea corectă a distribuției spațiale este necesară separarea *tendinței regionale* prezentă pe direcțiile în care structura are caracter *nestaționar* de cea *aleatoare* prezentă în cazul structurilor *staționare*.

COMENTARIU

Analiza staționarității ca și cea parametrică ne obligă să utilizăm instrumente geostatistice valabile numai în condiții strict definite.

De cele mai multe ori datele disponibile nu au caracteristicile adecvate prelucrării cu aceste instrumente și pentru a beneficia de numeroasele avantaje ale utilizării lor trebuie să facem anumite transformări sau concesiuni privind condițiile de utilizare.

Cât de mică trebuie să fie variația valorii variogramei în raport cu distanța pentru a accepta că o structură spațială este staționară?

Cât de mare trebuie să fie variația valorii variogramei pentru a face necesară eliminarea caracterului nestaționar al structurii?

Limitele acceptate sunt legate de concesiunile făcute relativ la condițiile de utilizare a instrumentelor geostatistice:

- *Utilizate în condițiile pentru care au fost definite, instrumentele geostatistice conduc la rezultate ce pot fi interpretate cu exactitate, eliminându-se ambiguitățile.*
- *Aplicate în cu totul alte condiții decât în cele pentru care au fost definite, instrumentele geostatistice conduc la rezultate neinterpretabile.*
- *Între aceste două extreme există situații intermediare în care de cele mai multe ori ne aflăm datorită extraordinarei variabilități a parametrilor ambientali.*

Pentru situațiile intermediare, în care ne asumăm riscul unor abateri de la condițiile stricte, efectul îl vom resimți în etapele ulterioare de prelucrare: estimarea distribuției spațiale și calculul erorilor. În etapa de testare a corectitudinii rezultatelor estimării distribuției spațiale se cumulează toate abaterile comise pe parcursul etapelor anterioare de prelucrare.

Când se evaluează eroarea de estimare nu se va putea ști însă cât din eroarea comisă este datorată nerespectării criteriului staționarității structurii pentru că eroarea poate fi determinată și de neluarea în considerare a anizotropiei sau de alegerea unui model neadecvat pentru variogramă.

Obstacolele din calea încercării de separare a efectelor diferitelor cauze care produc eroare de estimare sunt numeroase, motiv pentru care se recomandă reducerea la maximum a concesiunilor în aplicarea corectă a instrumentelor geostatistice. Un pas esențial pe calea respectării acestei tactici este identificarea nestaționarității structurilor analizate.

Parcursul superficială a acestei etape conduce la ignorarea nestaționarității structurii și la introducerea unor erori sistematice de estimare a căror evaluare este imposibilă.

Atenție !

După ignorarea analizei normalității distribuției frecvențelor valorilor, neglijarea analizei staționarității este cea mai frecventă eroare metodologică compromițând și ea operația de estimare a distribuției spațiale.

A13. Eliminarea nestaționarității

Să se analizeze staționaritatea distribuției sarcinii piezometrice a unui acvifer sub presiune explorat prin 24 de piezometre (**Fig.1.40**) în care s-au măsurat cotele nivelului piezometric (**tabelul 1.19**).

Rezolvare:

Tabelul 1.19 Coordonatele forajelor de explorare și cotele nivelului piezometric.

Nr.	x[km]	y[km]	NP[m]
1	13.46	76.22	169.66
2	8.19	24.76	135.91
3	24.11	47.53	156.78
4	40.97	53.27	168.94
5	64.45	13.24	164.18
6	16.95	50.83	157.96
7	85.91	39.19	191.04
8	1.71	97.83	166.56
9	23.73	6.49	142.66
10	7.19	45.66	148.71
11	29.32	86.62	175.68
12	6.89	39.44	146.84
13	48.61	53.41	174.11
14	87.13	85.96	212.05
15	51.38	91.78	191.36
16	58.13	9.14	161.65
17	15.13	46.58	154.79
18	22.78	16.13	144.13
19	51.92	47.49	172.47
20	44.27	78.99	182.95
21	76.81	29.64	184.89
22	57.31	0.76	154.10
23	0.65	32.79	135.53
24	71.62	93.68	206.41

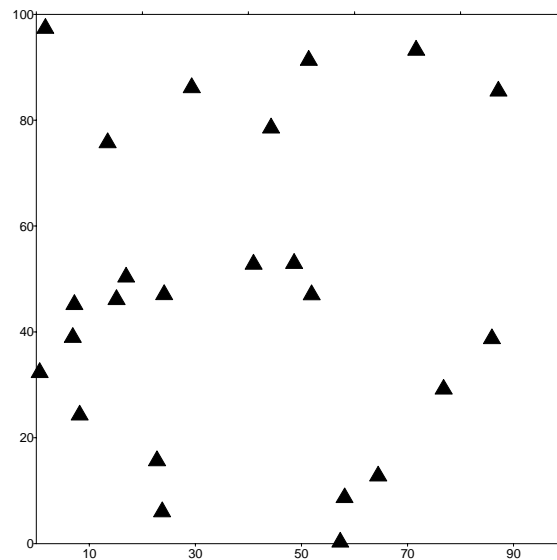


Fig.1.40 Forajele de explorare ale acviferului

Instrumentul complet pentru identificarea caracterului staționar sau nestaționar al suprafeței piezometrice este *variograma de suprafață* (**Fig.1.41**). Pentru suprafața piezometrică a acviferului studiat se remarcă variabilitatea nulă a sarcinii piezometrice pe direcția N45°V și nestaționaritatea ei pe direcția N45°E.

Pe direcția N45°E, variograma (**Fig.1.42**) are o creștere continuă, dovadă a existenței unei componente regionale cu *tendință* de creștere de la SV spre NE, pe care intenționăm în continuare să o calculăm și să o separăm de componenta aleatoare a suprafeței piezometrice a acviferului.

Pentru suprafața de tendință *alegem* o ecuație de gradul întâi de forma:

$$NP_{tend} = A + B \cdot x + C \cdot y$$

ai cărei coeficienți sunt soluțiile sistemului:

$$\begin{cases} 24 \cdot A + 908,61 \cdot B + 1167,45 \cdot C = 3999,35 \\ 908,61 \cdot A + 51544,61 \cdot B + 44518,05 \cdot C = 161419,05 \\ 1167,45 \cdot A + 44518,04 \cdot B + 76975,66 \cdot C = 203581,78 \end{cases}$$

Soluțiile sistemului ai cărui termeni sunt calculați în **tabelul 1.20** sunt : A=120, B = 0,50 și C = 0,45, ecuația suprafeței de tendință devenind:

$$NP_{tend} = 120 + 0,5 \cdot x + 0,45 \cdot y .$$

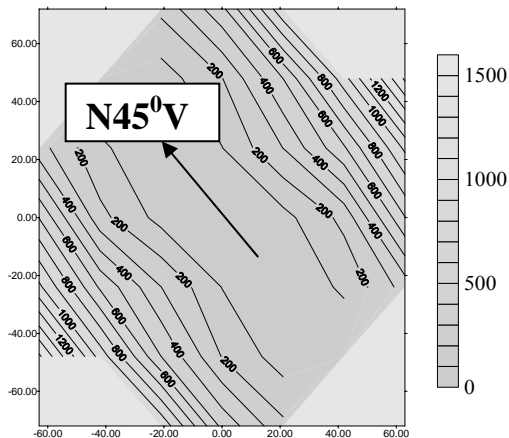


Fig.1.41 Variograma de suprafață a suprafeței piezometrice a acviferului sub presiune

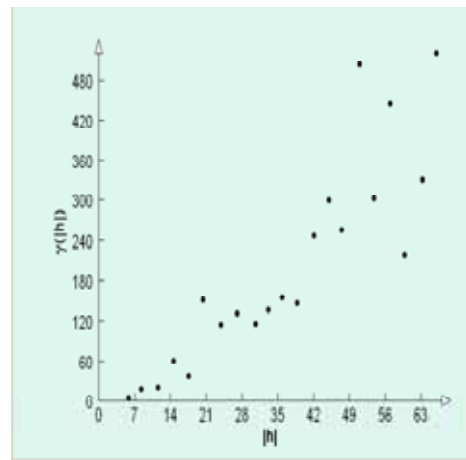


Fig.1.42 Variograma pe direcția N45°E

Puterea de ajustare a acestei suprafețe de tendință (P_{ε}) calculată pe baza valorilor din **tabelul 1.20** este:

$$P_{\varepsilon} = 100 \left[1 - \frac{501,48}{9432,56} \right] = 94,68 ,$$

ceea ce arată că variația spațială a sarcinii piezometrice este reprezentată în principal de funcția de tendință aleasă. Abaterile de la această tendință pot fi puse pe seama neomogenităților locale din acvifer. După eliminarea tendinței, valoarea rămasă a sarcinii piezometrice ($\Delta NP_i = NP_{mas_i} - NP_{tend_i}$) are o distribuție strict aleatoare, lucru ilustrat de variograma tip “efect de pepită total” obținută cu aceste valori (**Fig.1.43**).

Lipsa totală de structurare spațială a rezidului sarcinii piezometrice este justificată de puterea mare de ajustare a suprafeței de tendință de gradul întâi calculate. Mai sugestivă pentru semnificația componentelor structurii este reprezentarea prin hărți conturale a celor trei componente ale sarcinii piezometrice cu care s-a operat:

- sarcina piezometrică măsurată ($NP_{m\grave{a}s}$) (**Fig.1.44a**);
- sarcina piezometrică de tendință (NP_{tend}) (**Fig.1.44b**);
- reziduul sarcinii piezometrice (ΔNP) (**Fig.1.44c**).

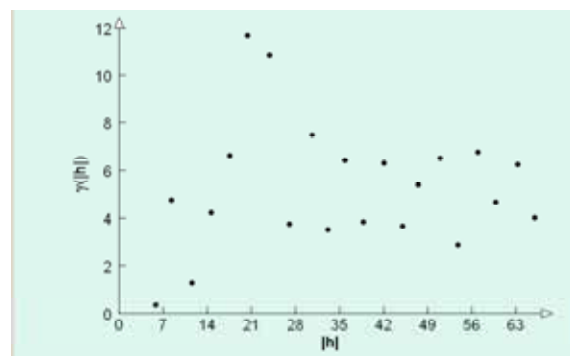


Fig.1.43 Variograma valorilor sarcinii piezometrice rămase după eliminarea componentei de tendință de gradul unu.

Din compararea celor trei hărți se remarcă amplitudinea redusă a abaterilor locale (exprimate de ΔNP) și similaritatea hărții piezometrice măsurate cu cea a hărții piezometrice de tendință.

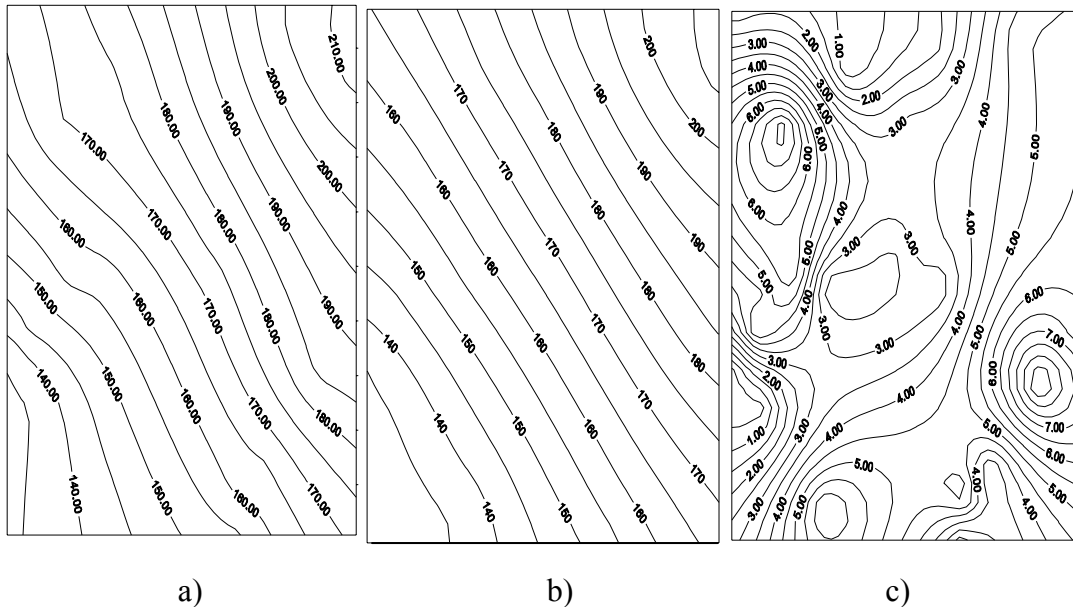


Fig.1.44. Hărțile conturale pentru “componentele” suprafeței piezometrice:
 a) sarcina piezometrică măsurată ($NP_{m\acute{a}s}$);
 b) sarcina piezometrică de tendință (NP_{tend});
 c) reziduuul sarcinii piezometrice (ΔNP)

Sunt situații în care componenta reziduală este semnificativă și conține elemente de structură spațială. În acele situații componenta reziduală se prelucrează separat de cea de tendință, prin *kriging ordinar*, iar în final se însumează cu aceasta.

COMENTARIU

Dacă structura spațială este **staționară**, adică toate valorile din spațiul cercetat oscilează în jurul unei **valori medii**, **aceeași pentru orice zonă a domeniului**, trecem la etapa următoare de prelucrare a datelor disponibile cu siguranța că estimările nu vor fi afectate de **erori sistematice**.

Dacă am identificat prezența nestaționarității este necesar să găsim **forma analitică adecvată** pentru aceasta. Procedeele de estimare a distribuției spațiale diferă pentru cele două componente ale valorilor măsurate: componenta de **tendință** (metode deterministe) și cea **aleatoare** (metode probabiliste).

Există metode geostatistice pentru prelucrarea simultană a celor două componente fără a introduce erori sistematice (**kriging universal**). Pentru aplicarea acestor metode este necesară cunoașterea **forme analitice a componentei de tendință**.

Nu lăsați “tendințele” să afecteze realismul estimărilor !!!

1.3.4. ANALIZA VARIOGRAFICĂ

Analiza structurală a unui fenomen regionalizat are ca obiectiv găsirea unui *model* al structurii. Elaborarea modelului face apel la cunoașterea fenomenului fizic studiat și la experiența în domeniul ajustării modelelor topo-probabiliste.

Instrumentul utilizat pentru identificarea modelului structural este *variograma*, motiv pentru care analiza structurală este cunoscută și sub denumirea de *analiză variografică*. *Variograma* este utilizată deoarece elimină calculul *mediei* valorilor, parametru cu semnificație ambiguă în cazul variabilelor nestaționare.

Obiectivul analizei variografice fiind în esență descriptiv, nu există constrângeri teoretice și în consecință orice tip de prelucrare este acceptată dacă reușește să clarifice corelația între *distanță* și *varianța erorii de estimare*.

Proprietățile variogramei care o fac instrumentul ideal pentru identificarea modelului structurii spațiale sunt sintetizate în:

$$\begin{cases} \gamma(0) = 0 \\ \gamma(h) = \gamma(-h) \geq 0 \end{cases} \quad (1.43)$$

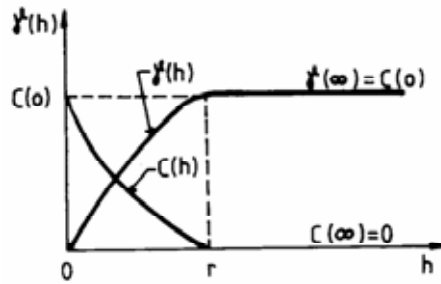


Fig.1.45. Variograma și covarianța.

Variograma ($\gamma(h)$; **Fig.1.45**) crește proporțional cu h . Când structura spațială pentru care este calculată variograma are caracter *staționar* valoarea maximă a acesteia rămâne constantă pentru valori ale lui h superioare unei anumite valori r , numită *rază de influență*. Se demonstrează că această valoare maximă numită *palier* nu este altceva decât varianța funcției aleatoare:

$$\gamma(\infty) = \text{Var}\{V(p)\} = c(0) \quad (1.44)$$

Dacă acest palier există, rezultă că și *covarianța complementară* există:

$$c(h) = c(0) - \gamma(h) \quad (1.45)$$

O variogramă cu *palier* și *rază de influență* caracterizează un fenomen regionalizat ce poate fi generat de o funcție aleatoare staționară de ordinul al doilea.

Zona de influență a informației dintr-un punct ($v(p)$) este dată de mărimea razei de influență r . În afara acestei zone, pentru $h > r$, nu mai există corelații între valorile variabilei măsurate: $v(p)$ și $v(p+h)$, ($c(h) = 0$ pentru $h > r$).

Într-un spațiu bidimensional sau tridimensional h reprezintă un *vector*. Majoritatea structurilor spațiale fiind *anizotrope*, valorile maxime ale variogramei și razele de influență sunt diferite, funcție de direcția de calcul.

Lipsa palierului pentru variogramă indică *nestaționaritatea* fenomenului regionalizat și necesitatea eliminării tendințelor structurate regional, pentru adecvarea semnalului prelucrat, modelului *funcției aleatoare*.

Analiza omnidirecțională este primul pas al identificării modelului structural și dă răspuns la întrebarea:

Există un model structural pentru variabila studiată ?

altfel spus:

Există o lege de variație spațială care să permită estimarea valorii variabilei studiate în orice punct al domeniului spațial cercetat?

Pentru a răspunde afirmativ sau negativ la această întrebare este necesar calculul *variogramei omnidirecționale* pentru care *toleranța direcțională* ($\Delta\theta$; **Fig.1.46**) este $\Delta\theta = 90^\circ$, suficient de mare astfel încât orientarea vectorului separator h_{ij} să nu mai influențeze valoarea variogramei.

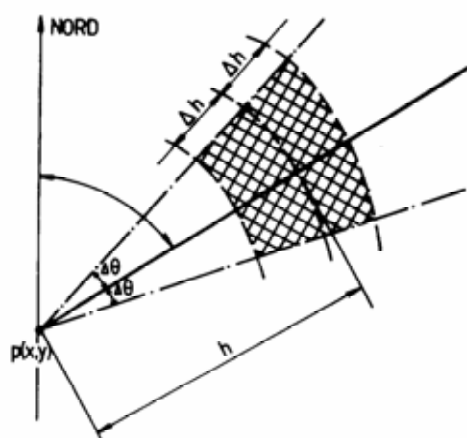


Fig.1.46.Elementele de toleranță utilizate pentru selecția perechilor de valori separate de vectorul h ($N(h)$): $\Delta\theta$ -toleranța de orientare; Δh -toleranța de distanță).

Variograma omnidirecțională poate fi considerată ca o *medie* a diverselor *variograme direcționale*. Dacă locațiile punctelor de observație sunt amplasate preferențial pe o anumită direcție, variograma omnidirecțională nu mai poate fi considerată ca o medie a variogramelor direcționale ea reflectând în principal particularitățile continuității pe acea direcție. *Variograma omnidirecțională* conține maximum de perechi de puncte în raport cu orice variogramă direcțională și dacă ea nu reflectă existența unui model structural, nici o variogramă direcțională nu va avea o șansă mai mare de reușită.

Calculul variogramei omnidirecționale nu implică ipoteza izotropiei structurii, ci o primă etapă a identificării modelului structural, utilă stabilirii corecte a *parametrilor de distanță* necesari identificării variogramei reprezentative a structurii.

Parametrii de distanță care condiționează în mod direct valorile variogramelor omnidirecționale dar și unidirecționale sunt:

- *distanțele* pentru care se calculează variograma (h_i ; $i=0,1,2,\dots$);
- *toleranța de distanță* (Δh).

Dacă punctele de observație sunt plasate într-o *rețea regulată*, *distanțele* pentru care se calculează variograma (h_i) sunt un multiplu al parametrului rețelei de probare, distanța inițială fiind chiar acest parametru. În cazul unei *rețele neregulate*, *distanța minimă* pentru care se calculează variograma este o medie a distanței dintre punctele vecine, celelalte distanțe fiind un multiplu al acestei distanțe. *Valoarea maximă a distanței* pentru care se calculează variograma se limitează la jumătate din distanța maximă dintre punctele de observație disponibile. Alegerea finală a distanțelor de calcul este determinată de aspectul funcției de variogramă. *Clarificarea* variogramei impune modificarea distanțelor inițiale pe baza efectului lor. În general, creșterea distanțelor de calcul determină o netezire a variogramei experimentale.

Toleranța de distanță (Δh , **Fig.1.46**) este introdusă în calcul pentru a crește numărul de perechi de valori corespunzătoare unei anumite clase de distanță într-o rețea neregulată de probare, fiind puțin probabil ca mai multe perechi de valori să fie plasate riguros la aceeași distanță. Toleranța de distanță se alege în general ca jumătate din decalajul de distanță dintre două clase consecutive, iar din necesități de detaliere poate fi redusă. Utilizarea unui decalaj constant între clase și a unei toleranțe egale cu jumătate din acest decalaj asigură acoperirea completă a suprafeței studiate și utilizarea tuturor valorilor disponibile din rețeaua de probare.

Analiza unidirecțională demarează după ce s-a obținut o variogramă omnidirecțională clară ce indică existența unui model structural și răspunde la întrebarea:

Modelul structural pentru variabila studiată este același pentru orice direcție ?

sau altfel spus:

Modelul structural este izotrop sau anizotrop?

Există numeroase informații utile identificării *direcțiilor de anizotropie* ale unui model structural. Pentru studiul distribuției poluanților în aer, aceste informații provin din cunoașterea direcției vânturilor; în studiul transportului poluanților din apa subterană provin din caracteristicile structurale și dinamice ale acviferului iar în studiul concentrațiilor minerale în depozitele stratiforme din orientarea stratificației.

Calea cea mai riguroasă de identificare a direcțiilor de anizotropie este calculul *variogramelor direcționale* cu ajutorul cărora se construiesc: *variograma de suprafață* sau *diagrama razelor de influență ale variogramei*.

Pentru calculul *variogramelor direcționale* se alege o *toleranță unghiulară minimă* ($\Delta\theta$) pentru a elimina pe cât posibil mascarea anizotropiei. Din păcate, cu o toleranță direcțională prea mică se obțin puține perechi de puncte pentru fiecare clasă de distanță, iar variogramele direcționale sunt inutilizabile pentru identificarea modelului structural. Identificarea toleranței direcționale optime se face prin încercări experimentale bazate pe compararea variogramelor direcționale calculate pentru diferite toleranțe.

Analiza completă a anizotropiei se bazează pe *variograma de suprafață* care este o hartă conturală a variogramelor direcționale. Deoarece programele automate pentru trasarea hărților conturale lucrează cu o rețea rectangulară de puncte, este eficient, din punct de vedere al calculelor, să se utilizeze pentru fiecare clasă de distanță o toleranță definită într-un sistem rectangular de coordonate. Pentru calculul valorii variogramei direcționale a perechilor de valori separate de vectorul $h=(h_x, h_y)$ se grupează toate punctele care pe direcția Ox sunt separate de distanța $h_x+\Delta x$ iar pe

direcția axei OY de distanța $h_y + \Delta y$ (**Fig.1.47**). În această variantă, *sectorul de coroană circulară* utilizat (în mod clasic) pentru selectarea punctelor dintr-o clasă de distanță (**Fig.1.46**), se transformă într-un *dreptunghi* (**Fig.1.47**).

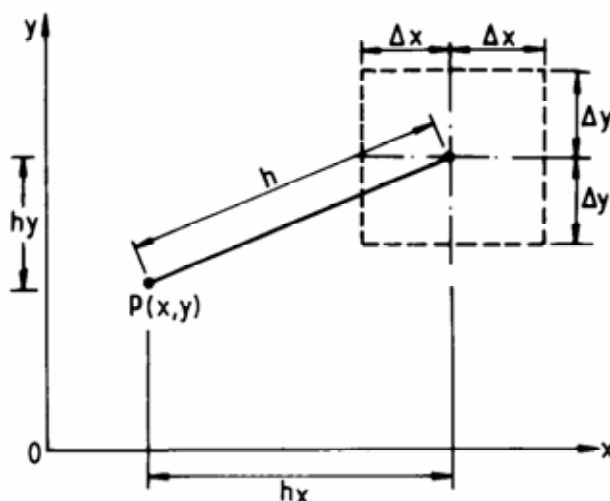


Fig.1.47.Elementele de calcul pentru variograma de suprafață

Când dispunem de un număr mic de puncte de observație, pe baza unui număr redus de variograme direcționale se realizează *diagrama razelor de influență* ale variogramei. *Diagrama razelor de influență* permite evaluarea raportului de anizotropie al modelului structural. *Raportul de anizotropie* (R_a) este exprimat cantitativ prin raportul dintre raza de influență pe direcția de variabilitate minimă (R_m) și cea pe direcția de variabilitate maximă (R_M) ($R_a = R_m/R_M$).

Variogramele relative sunt utilizate atunci când rezultatul analizei omnidirecționale este negativ.

Dacă *variograma omnidirecțională* nu reflectă o corelație între distanța (h) și pătratul creșterii medii ($\gamma(h)$), examinarea *diagramelor de continuitate* este obligatorie pentru identificarea punctelor care afectează această corelație. Rolul unor perechi de puncte sau a unor puncte izolate poate fi semnificativ și eliminarea lor poate îmbunătăți gradul de corelație. Dacă încercările de îmbunătățire a acestei corelații prin eliminarea punctelor nu conduce la nici un rezultat, în această etapă descriptivă de analiză a continuității pot fi utilizate și alte funcții de distanță de tipul *variogramelor relative*.

Dependența variogramei față de media valorilor din fiecare clasă de distanță determină utilizarea *variogramelor relative*. Variogramele relative se calculează prin raportarea valorii variogramei la diferite tipuri de medii locale. Ele sunt utilizate când variogramele absolute nu indică o lege de variație spațială a variabilei studiate.

În mod curent se utilizează trei tipuri de variograme relative: *variograma relativă locală* (γ_{RL}), *variograma relativă generală* (γ_{RG}) și *variograma relativă pereche* (γ_{RP}).

Variograma relativă locală este recomandată atunci când legea de distribuție a variabilei studiate este multimodală. Precedată de o analiză dispersională care să permită separarea domeniului spațial în subzone omogene (Scrădeanu D.,1995) variograma relativă locală se calculează cu relația:

$$\gamma_{RL}(h) = \frac{\sum_{i=1}^r N_i(h) \cdot \frac{\gamma_i(h)}{m_i^2}}{\sum_{i=1}^r N_i(h)} \quad (1.46)$$

în care:

$\gamma_i(h)$ - variogramele locale calculate pentru fiecare din cele r regiuni separate
($i=1, \dots, r$) (**Fig.1.48**);

m_i - mediile locale pentru cele r regiuni separate ($i = 1, \dots, r$);

$N_i(h)$ - numărul perechilor de valori pentru fiecare regiune.

Ecuția de calcul raportează variogramele locale la pătratul mediilor locale prin luarea în considerare și a numărului de perechi de valori pe care se bazează fiecare. Raportarea variogramei locale la pătratul mediei locale este valabilă numai în cazul unei corelații lineare (efect de proporționalitate linear) între media locală și abaterea standard locală. Pentru un efect de proporționalitate nelinear, raportarea variogramelor locale se face la alte funcții de media locală.

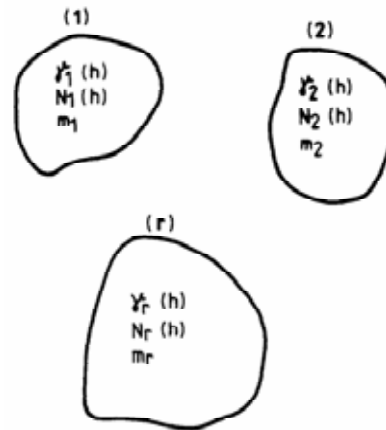


Fig.1.48.Elementele de calcul pentru variograma relativă locală.

Variograma relativă generală compensează unul din neajunsurile esențiale ale variogramei relative locale și anume numărul redus de puncte pe care se bazează unele dintre componente datorită separării suprafeței totale în subzone cu extindere redusă. Utilizarea unui număr redus de puncte poate conduce la o variogramă fluctuantă, imposibil de utilizat pentru analiza continuității și estimărilor spațiale care succed acesteia.

Variograma relativă generală nu necesită separarea unor subzone. Ea se calculează raportând suma pătratelor diferențelor dintre perechile de valori, separate de distanța h , la pătratul mediilor aritmetice:

$$\gamma_{RG}(h) = \frac{\gamma(h)}{m(h)^2} \quad (1.47)$$

Numărătorul este variograma experimentală absolută iar numitorul este pătratul mediei calculate cu relația:

$$m(h) = \frac{1}{2N(h)} \sum_{(i,j)|h_{ij}=h}^{N(h)} v_i + v_j = \frac{m_{+h} + m_{-h}}{2} \quad (1.48)$$

Calculul variogramei relative generale are ca efect reducerea fluctuațiilor funcției de variogramă.

Variograma relativă pereche ajustează variograma absolută în raport cu pătratul mediei fiecărei perechi de valori:

$$\gamma_{RR}(h) = \frac{1}{2N(h)} \sum_{(i,j)|h_{ij}=h}^{N(h)} \frac{(v_i - v_j)^2}{\left(\frac{v_i + v_j}{2}\right)^2} \quad (1.49)$$

Variograma relativă pereche este utilă, în cazul existenței unor valori extreme, în selecția de date disponibile, ea permițând reducerea influenței acestora asupra valorilor variogramei.

Din punct de vedere operațional, calculul *variogramei relative pereche* impune introducerea unei valori minime a sumei celor două valori v_i și v_j deoarece atunci când aceasta este zero valoarea variogramei relative pereche este nedefinită. Pentru valori ale sumei $v_i + v_j$ mai mici decât această limită, perechile de valori se exclud din calcul.

Utilizarea variogramelor relative conduce la eliminarea fluctuațiilor *variogramei absolute* atât pentru valori mici ale distanței h , prin utilizarea *variogramei relative locale*, cât și pentru valori mari ale distanței h , prin utilizarea *variogramei relative generale* și *variogramei relative pereche*.

Interpretarea variogramelor experimentale se bazează pe graficul acestora și are ca obiectiv stabilirea strategiei de modelare matematică.

Forma graficului variogramei experimentale exprimă sintetic caracteristicile structurii iar analiza ei permite separarea acestora în două categorii:

- caracteristici structurale locale (*continuitate și regularitate*), rezultate din forma variogramei pentru distanțe (h) mici (*în vecinătatea originii*);
- caracteristici structurale regionale (*stiluri structurale*), rezultate din forma variogramei pentru distanțe (h) mari (*la infinit*).

Continuitatea și regularitatea în spațiu a variabilelor regionalizate sunt legate de forma graficului variogramei în vecinătatea originii. În ordinea descreșterii regularității se disting patru tipuri de forme ale variogramei în vecinătatea originii:

- forma parabolică (**Fig.1.49a**):

$$\gamma(h) \sim \lambda|h|^2, \quad h \rightarrow 0$$

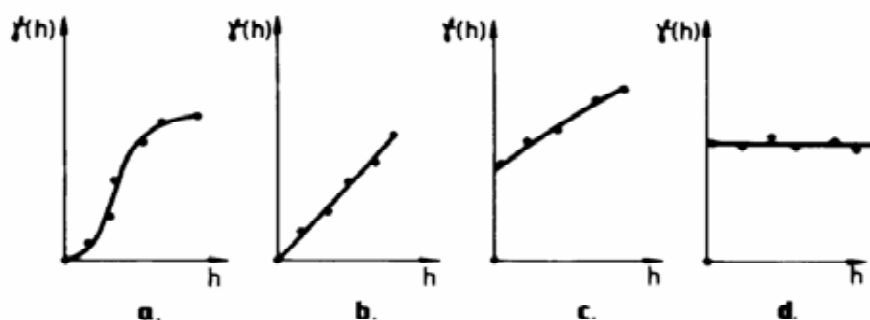


Fig.1.49. Formele variogramei în vecinătatea originii.

Funcția de variogramă este în acest caz de două ori derivabilă în origine și deci funcția aleatoare asociată variabilei regionalizate are și ea derivata întâi. Această formă indică o foarte bună continuitate și regularitate a variabilei regionalizate.

- forma lineară fără efect de pepită (**Fig.1.49b**):

$$\gamma(h) \sim \lambda|h|, h \rightarrow 0 \quad (1.50)$$

Funcția de variogramă nu mai este derivabilă în origine dar rămâne continuă pentru $h = 0$. Funcția aleatoare asociată variabilei regionalizate nu este derivabilă și indică o regularitate mai redusă.

- discontinuitate în origine (**Fig.1.49c**):

$$\lim_{h \rightarrow \infty} \gamma(h) \neq 0 \quad (1.51)$$

Chiar dacă, prin definiție, variograma este nulă în origine, în anumite situații variabilitatea între două valori foarte apropiate poate fi mare. La această variabilitate locală, comparabilă cu zgomotul de fond al unei anomalii, se adaugă, pentru distanțe mai mari, o variabilitate continuă tradusă prin continuitatea variogramei pentru $h > 0$. Această discontinuitate în origine, cunoscută sub numele de *efect de pepită*, poate fi datorată erorilor de măsură sau microvariabilității nedecelabile la densitatea informației disponibile.

- efect de pepită pur (**Fig.1.49d**):

$$\gamma_0(h) = \begin{cases} 0, & h = 0 \\ 1, & h > 0 \end{cases} \quad (1.52)$$

Este cazul limită al situației precedente când pentru $h > 0$ variograma nu mai are continuitate. Modelarea unei astfel de variograme se realizează printr-un model cu palier cu raza de influență infinit mică. Prezența *efectului de pepită total* indică absența oricărei corelații spațiale.

Intuirea semnificației acestor condiții de natură matematică, proprii instrumentelor utilizate, este premiza unei corecte manevrări a acestora pentru cuantificarea aspectelor calitative semnalate în etapa de reprezentare grafică a datelor disponibile. Doar corelarea rezultatelor prelucrărilor cu fenomene naturale expresive poate fundamenta însușirea corectă a tehnicilor geostatistice și alegerea corectă a variantelor de lucru.

Un exemplu clasic pentru înțelegerea semnificației comportamentului variogramei în vecinătatea originii este examinarea variogramelor experimentale ale nivelului piezometric din acviferul Korhongo (Coasta de Fildeș; J.P.Delhomme, 1977).

Datele prelucrate sunt adâncimi ale nivelului piezometric $v(x,t)$ măsurate în timp în patru piezometre amplasate la 500 m unul de altul (**Fig.1.50**).

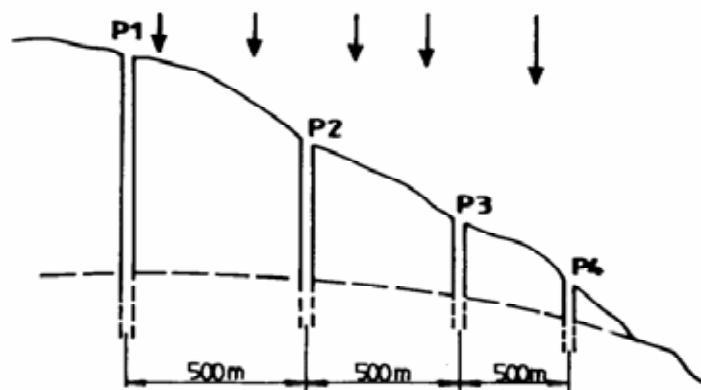


Fig.1.50. Piezometre în acviferul Korhongo (Coasta de Fildeș; J.P.Delhomme, 1977)

Evoluția nivelului piezometric în cele patru piezometre indică o variabilitate diferită determinată de grosimea zonei de aerare cu reflectare elocventă în forma variogramelor (Fig.1.51):

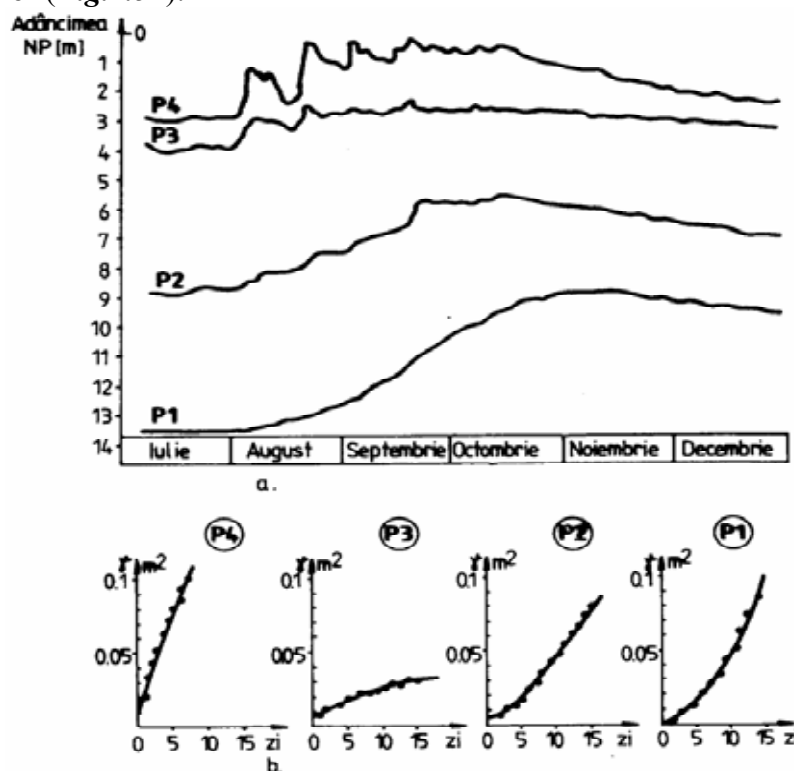


Fig.1.51. Analiza variografică a nivelului piezometric din acviferul Korhongo : a - evoluția profilurilor piezometrice ; b - variogramele evoluției în timp a nivelurilor piezometrice.

- în P_4 nivelul piezometric este foarte aproape de suprafața solului și reacționează la fiecare aversă. Profilul piezometric este neregulat (Fig.1.51a) iar variograma corespunzătoare este lineară, cu efect de pepită (Fig. 1.51b);

- în P_3 zona de aerare este mai grosă și amortizează efectul precipitațiilor asupra variațiilor piezometrice. Variograma corespunzătoare prezintă un efect de pepită mai mic și pentru $h > 0$ variograma are o formă curbilinie ce indică o trecere spre un model cu plafon;

- în P_2 adâncimea nivelului piezometric este mare, astfel încât efectul de amortizare al variațiilor alimentării prin infiltrații este aproape total, resimțindu-se

numai efectul unor alimentări bruște. Variograma are un efect de pepită redus ce reflectă prezența acestor alimentări bruște iar pentru $h > 0$ un comportament parabolic ce indică o bună continuitate;

- în P_1 nu se mai resimte influența alimentării discontinui a acviferului, profilul piezometric prezentând o variație continuă. Variograma corespunzătoare este de tip parabolic, fără efect de pepită.

Stilurile structurale se identifică pentru valori mari ale distanței h la care există patru tipuri de funcții de variogramă care implică interpretări și strategii operaționale distincte:

- *creșterea continuă a variogramei*, proporțional cu distanța dintre puncte, reflectă absența staționarității structurii.

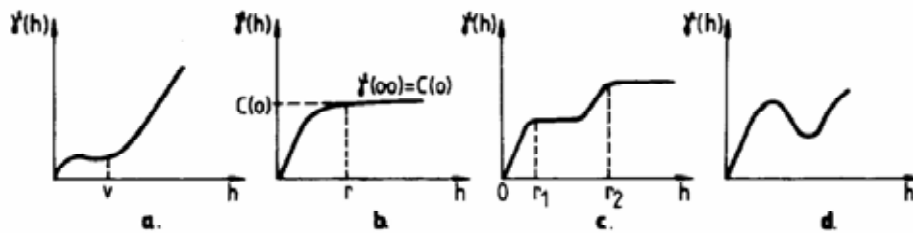


Fig. 1.52. Comportamentul variogramei la infinit.

Tendința care induce nestaționaritatea este frecvent legată de o anizotropie geometrică, variograma comportându-se de manieră diferită, funcție de direcție. Dacă scara de lucru este relativ mare, variograma experimentală poate lua o alură parabolică de la o anumită distanță $h > r$ (**Fig. 1.52a**). Pentru distanțe inferioare acesteia (interpolări în vecinătăți ce nu depășesc distanța r), ipoteza de staționaritate, în sens larg, poate fi adoptată iar influența derivatei este neglijabilă dacă în această vecinătate variograma este izotropă și nu depinde de punctul de aplicație.

- *atingerea unei valori maxime a variogramei* (palierul variogramei), pentru un h finit (raza de influență a variogramei), indică prezența unei structuri simple staționare a variabilei regionalizate (**Fig. 1.52b**);

- *stabilizarea valorii variogramei* pe două sau mai multe paliere semnaleză prezența unor structuri staționare suprapuse caracterizate prin variograme cu modele diferite și parametri diferiți (**Fig. 1.52c**). Deconvoluția variogramei în componentele elementare chiar dacă nu este unică și necesită informații complementare permite separarea structurilor suprapuse.

- *descreșterea temporară a variogramei* pentru anumite distanțe (**Fig. 1.52d**) este dificil de interpretat, ea putând fi determinată de repartiția neuniformă a punctelor de observație sau de existența unei componente periodice a structurii. Eliminarea acestor neregularități se realizează prin intermediul variogramei relative locale sau se modelează cu ajutorul modelelor trigonometrice.

Modelarea variogramelor experimentale este necesară deoarece toate operațiunile de estimare de tip topo-probabilist se bazează pe variogramă și acestea nu pot fi efectuate cu o funcție tabelară așa cum este variograma experimentală. Realizarea calculului de estimare implică identificarea unui model teoretic care să interpozeze cât mai bine valorile variogramei experimentale.

În alegerea *modelelor teoretice*, două caracteristici ale variogramei experimentale sunt luate în considerare: comportarea în vecinătatea originii (parabolică, lineară sau efect de pepită total) și prezența sau absența palierului.

Funcție de aceste caracteristici modelele teoretice utilizate în mod curent pot fi separate în trei categorii:

- *modele cu palier* (modele tranzitive) cu comportament linear în vecinătatea originii (*modelul sferic și modelul exponențial*) și cu comportament parabolic în vecinătatea originii (*modelul gaussian*);

- *modele fără palier* (modelul putere și modelul logaritmic);

- *modelul efect de pepită*.

Modelele teoretice cu care se operează sunt normate, adică au varianța unitară. Pentru a obține un model cu palier diferit de unu este suficient să se multiplice expresia normată a modelului cu o constantă.

Modelul sferic, cel mai frecvent utilizat, are ecuația:

$$\gamma(h) = \begin{cases} \frac{3}{2} \cdot \frac{h}{r} - \frac{1}{2} \cdot \frac{h^3}{r^3}, \forall h \in [0, r] \\ 1, \forall h \in [0, r] \end{cases} \quad (1.53)$$

Modelul exponențial are ecuația:

$$\gamma(h) = 1 - \text{EXP}\left(-\frac{h}{r}\right) \quad (1.54)$$

De remarcat că *modelul sferic* atinge valoarea maximă (palierul) pentru o distanță finită $h = r$ în timp ce *modelul exponențial* tinde asimptotic la aceasta (teoretic la $h = \infty$) (**Fig.1.53**). În practică, pentru *modelul exponențial* se adoptă o rază de influență efectivă $r' = 3r$, pentru care $\gamma(r') = 0,951$.

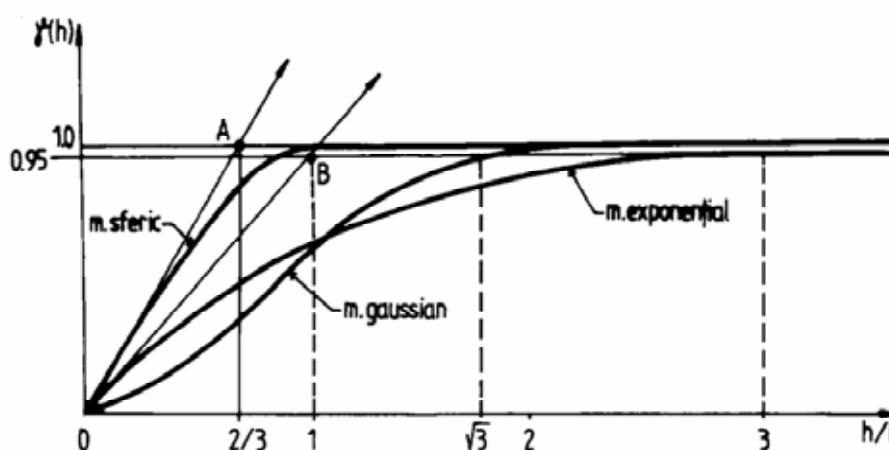


Fig.1.53. Modele cu palier (tranzitive)

Ceea ce diferențiază *modelul sferic* de cel *exponențial* sunt abscisele punctelor de intersecție dintre tangentele la origine și palier; pentru *modelul sferic*, două treimi din raza modelului ($x_A = 2r/3$), iar pentru *modelul exponențial* o treime din raza de influență efectivă ($x_B = r'/3$).

Modelul *gaussian* este valabil pentru un comportament foarte regulat în vecinătatea originii și are ecuația:

$$\gamma(h) = 1 - \text{EXP}\left(-\frac{h^2}{r^2}\right) \quad (1.55)$$

Palierul este atins asimptotic, raza de influență efectivă fiind $r' = r\sqrt{3}$ pentru o valoare a variogramei $\gamma(r') = 0,951 \approx 1,0$.

Comportamentul parabolic din vecinătatea originii al modelului gaussian nu trebuie confundat cu efectul unei derive deoarece la distanțe mai mari valoarea variogramei se stabilizează în jurul palierului. Pentru analize locale această confuzie este posibilă și ea trebuie evitată pentru corecta alegere a modelului.

Modele fără palier corespund variogramelor experimentale a căror creștere este continuă în limitele domeniului de observație.

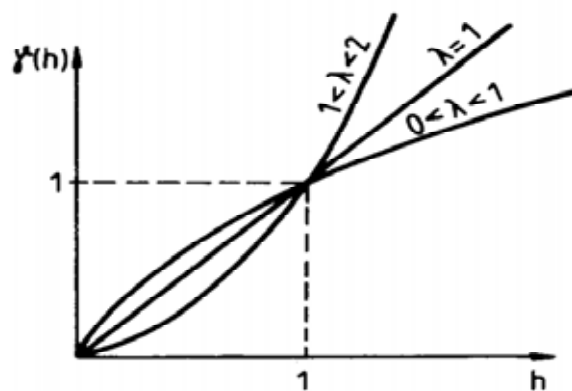


Fig. 1.54. Model putere

Modelul *putere* (**Fig. 1.54**) este de ecuație:

$$\gamma(h) = h^\lambda, \lambda \in (0, 2) \quad (1.56)$$

În practică, modelul linear ($\lambda=1$) este cel mai utilizat, el putând servi la estimarea în vecinătatea originii a tuturor modelelor cu comportament linear (modelul sferic și modelul exponențial). Pe măsură ce λ crește comportamentul în vecinătatea originii este mai regulat, dar pentru $\lambda > 2$ modelul nu mai este pozitiv definit fiind incompatibil cu evaluările geostatistice lineare.

Modelul *logaritm* (cunoscut și sub numele de modelul De Wijs) are ecuația:

$$\gamma(h) = \log h \quad (1.57)$$

Modelul logaritm nu poate descrie structurile de suport punctual, dar această condiție nu este deranjantă deoarece în practică datele experimentale care definesc variabila regionalizată sunt relative la un suport nepunctual (nenul), adică la o probă de o dimensiune finită.

Discontinuitatea în vecinătatea originii a variogramei (efectul de pepită) poate fi interpretată ca un model tranzitiv care atinge palierul la o rază de influență mai mică în raport cu distanțele dintre punctele de observație disponibile.

Modelele adecvate acestor situații poartă numele de *modele efect de pepită* și au ecuația:

$$\gamma(h) = \begin{cases} 0, & h = 0 \\ 1, & h > 0 \end{cases} \quad (1.58)$$

Modelul efect de pepită nu este considerat în mod uzual un model elementar dar apare ca o constantă în ecuația majorității modelelor de variogramă de diferite tipuri.

În *mediile izotrope* modelul variogramei depinde numai de distanță și este independent de direcție. În astfel de situații este necesară modelarea variogramei numai pe o direcție și de cele mai multe ori se preferă modelarea variogramei omnidirecționale (cu toleranța direcțională $\Delta\theta = 90^\circ$).

Deși de cele mai multe ori se poate modela satisfăcător o variogramă experimentală cu ajutorul unui model *elementar*, pentru o calare mai riguroasă se preferă un model *complex* obținut prin combinarea lineară a mai multor modele elementare:

$$\gamma(h) = \sum_{i=1}^n |w_i| \cdot \gamma_i(h) \quad (1.59)$$

Pentru calarea unei combinații de modele elementare la o variogramă experimentală direcțională trebuie identificat în primul rând modelul care dă caracteristica esențială (cu palier, fără palier, etc.). Deseori este necesară combinarea modelelor elementare din categorii diferite. Pentru o variogramă experimentală care nu atinge un palier dar are o comportare parabolică în vecinătatea originii este necesară combinarea unui model gaussian cu altul linear.

Începătorii în analiza variografică sunt tentați să complice modelele pentru o calare foarte exactă a variogramelor experimentale care în etapa de estimare spațială nu se justifică. Principiul economiei în definirea modelelor este un bun ghid în modelarea variogramelor experimentale.

În selectarea caracteristicilor variogramelor experimentale ce trebuie modelate este înțelept de luat în considerare existența unei explicații fizice a caracteristicii respective. Dacă informațiile calitative asupra cauzelor fenomenului a cărui structură spațială este studiată explică sau confirmă o anumită caracteristică a variogramei experimentale, atunci este necesar ca aceasta să fie conținută în model, în caz contrar ea poate fi considerată ca rezultat al unui fapt accidental și ignorată în procesul de modelare.

După alegerea modelelor elementare, modelarea variogramei experimentale se transformă într-un simplu exercițiu de calare în care se determină parametrii modelelor teoretice.

Deseori variogramele experimentale direcționale evidențiază schimbări majore cu direcția ca efect al anizotropiei structurilor. Există două tipuri distincte de anizotropie care comportă modalități diferite în definirea modelului de variogramă:

- *anizotropia geometrică*, în care variogramele direcționale au același model și palier în toate direcțiile și doar razele de influență sunt diferite (**Fig.1.55a**). Variograma de suprafață (**Fig.1.55b**) evidențiază anomalii alungite de minim în suprafețe orizontale.

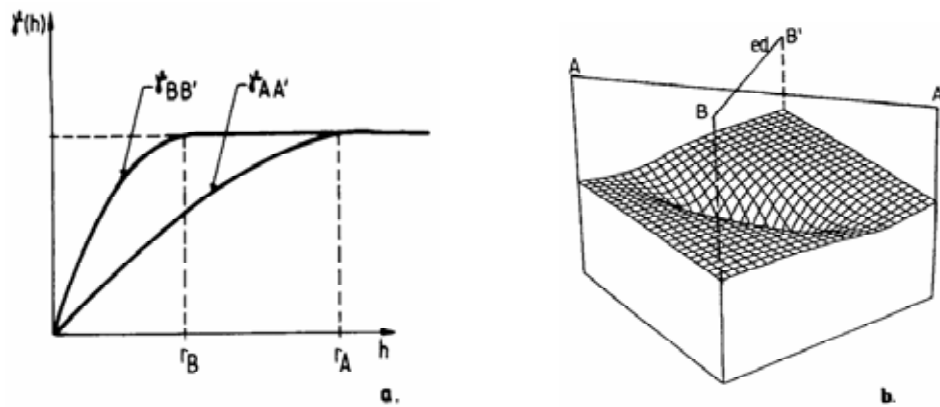


Fig.1.55. Anizotropie geometrică.

- *anizotropie zonală*, care se manifestă prin modificarea cu direcția a palierului și menținerea razei de influență și a modelului. Două variograme pe două direcții ortogonale (a) și variograma de suprafață (b) corespunzătoare pentru o astfel de anizotropie sunt prezentate în **Fig.1.56**.

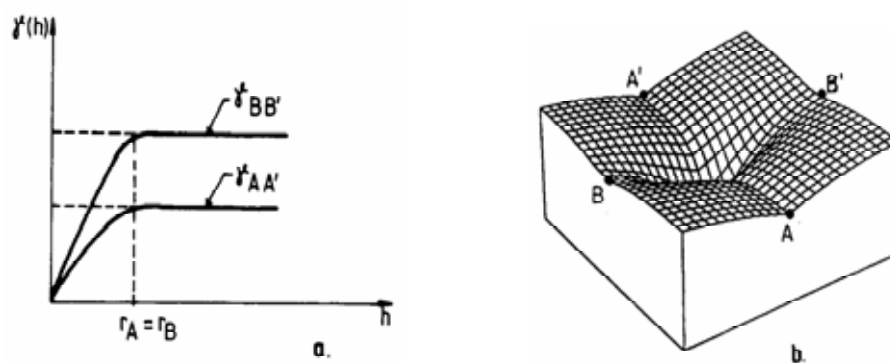


Fig.1.56. Anizotropie zonală

Identificarea axelor de anizotropie este prima operațiune care trebuie realizată în construirea modelelor anizotrope. Harta conturală a variogramei de suprafață și diagramele radiare sunt instrumentele utilizate în această operațiune.

După identificarea direcțiilor de anizotropie, pasul următor este construirea unui model (izotrop) echivalent care să permită cuantificarea modificării variogramei cu distanța și direcția. Modelul echivalent se elaborează într-o primă etapă în sistemul de referință al axelor de anizotropie și în etapa următoare în sistemul de referință al datelor primare.

Pentru definirea modelului echivalent diferitelor modele de variograme direcționale, se realizează o transformare care reduce toate variogramele direcționale la un model comun cu o rază standardizată unitară. Transformarea afectează doar distanța dintre perechile de valori, funcție de direcția pe care se calculează variograma. Modificarea distanței se realizează în așa fel încât valoarea variogramei calculată cu modelul echivalent pe baza distanței transformate (h_1) să fie identică cu cea calculată cu variograma direcțională pe baza distanței reale (h).

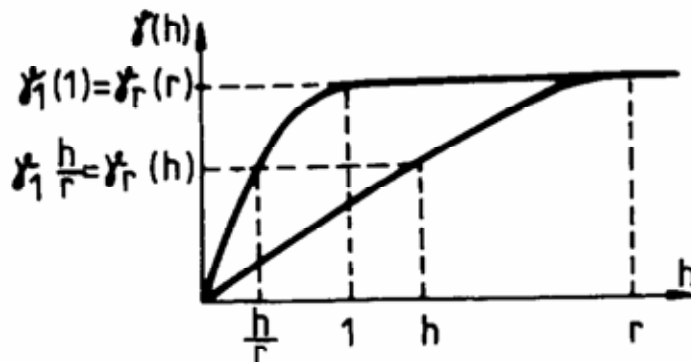


Fig.1.57. Model echivalent în sistemul de coordonate al axelor de anizotropie (anizotropie geometrică)

Pentru cazul unei anizotropii geometrice, descrise de două variograme direcționale cu razele de influență l și r (**Fig.1.57**), dacă evaluăm modelul cu rază unitară la o distanță h/r vom obține aceeași valoare calculată cu modelul cu rază r la distanța h . În acest mod se reduce modelul cu rază r la un model echivalent cu rază unitară prin reducerea distanței de la h la h/r . Această echivalență poate fi scrisă sub forma:

$$\gamma_1(h_1) = \gamma_r(h), \quad h_1 = \frac{h}{r} \quad (1.60)$$

Pentru un spațiu tridimensional modelul echivalent și distanța redusă sunt definite de relațiile:

$$\gamma(h) = \gamma(h'_x, h'_y, h'_z) = \gamma_1(h_1) \quad - \text{modelul echivalent} \quad (1.61)$$

$$h_1 = \sqrt{\left(\frac{h'_x}{r_x}\right)^2 + \left(\frac{h'_y}{r_y}\right)^2 + \left(\frac{h'_z}{r_z}\right)^2} \quad - \text{distanța redusă} \quad (1.62)$$

în care:

(h'_x, h'_y, h'_z) - proiecțiile vectorului separator h pe axele de anizotropie (a_x, a_y, a_z) (**Fig.1.58**);

α - unghiul de rotație în planul xOy ;

(r_x, r_y, r_z) - razele variogramelor direcționale pe direcțiile axelor de anizotropie (a_x, a_y, a_z) .

Pentru modelele fără plafon, în locul razelor r_x, r_y, r_z se utilizează pantele în origine ale variogramelor direcționale (i_x, i_y, i_z) , distanțele reduse corespunzătoare fiind $i_x \cdot h, i_y \cdot h$ și $i_z \cdot h$.

Dacă axele de anizotropie (a_x, a_y, a_z) nu coincid cu axele de coordonate în care sunt plasate datele primare (O_x, O_y, O_z) (**Fig.1.58**), modelul echivalent de variogramă trebuie transpus în sistemul de referință al datelor primare din considerente de eficiență a prelucrării.

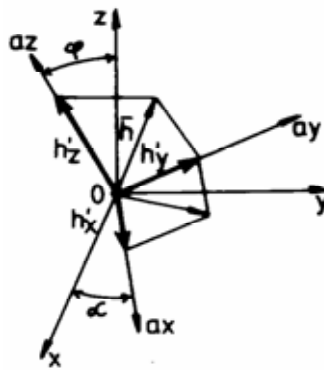


Fig.1.58. Componentele vectorului \bar{h} în sistemul de referință al axelor de anizotropie

Cea mai directă metodă pentru realizarea acestei transformări se bazează pe cunoașterea unghiului de rotație (α) în jurul axei Oz (necesar suprapunerii axelor a_x și a_y cu Ox , respectiv Oy) și a unghiului de rotație (φ) în jurul axei Oy (necesar suprapunerii axei a_z cu Oz).

Dacă notăm cu h' vectorul în sistemul de referință al axelor de anizotropie, cu h vectorul în sistemul de referință al datelor primare și cu R matricea de rotație, transformarea se realizează cu relația:

$$h' = R \cdot h \quad (1.63)$$

în care:

$$h' = [h'_x \ h'_y \ h'_z], \quad h = [h_x \ h_y \ h_z] \text{ și } R = \begin{bmatrix} \cos(\alpha) \cdot \cos(\varphi) & \sin(\alpha) \cdot \cos(\varphi) & \sin(\varphi) \\ -\sin(\alpha) & \cos(\alpha) & 0 \\ -\cos(\alpha) \cdot \sin(\varphi) & -\sin(\alpha) \cdot \sin(\varphi) & \cos(\varphi) \end{bmatrix} \quad (1.64)$$

După transformarea vectorului h din sistemul de referință al datelor primare în sistemul de referință al axelor de anizotropie, poate fi corect evaluat modelul anizotrop folosind vectorul h' .

Calculul vectorului h'_i conținând distanțele reduse poate fi combinat cu relația de transformare, rezultând într-o scriere compactă ecuația:

$$h'_i = T \cdot R \cdot h \quad (1.65)$$

Această ecuație exprimă ideea că vectorul de poziție h trebuie exprimat în sistemul de referință al axelor de anizotropie înaintea calculului distanțelor reduse.

Problema *modelării anizotropiei geometrice*, frecvent întâlnită în practică, este cea schematizată în **Fig.1.59**. Analiza se realizează într-un plan orizontal, fiecare din cele două modele direcționale calculate pe două direcții de anizotropie ortogonale fiind compus din trei modele elementare:

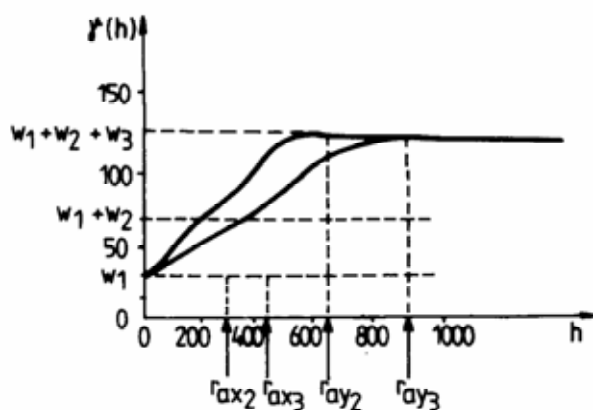


Fig.1.59. Model de anizotropie geometrică

- *primul model*: efect de pepită izotrop, dat de ecuația cu forma generală:

$$\gamma(h) = w_1 \cdot \gamma_1(h_1) \quad \text{și distanța redusă: } h_1 = h \quad (1.66)$$

- *al doilea model*: model de tip sferic, dat de ecuația cu forma

generală:

$$\gamma(h) = w_2 \cdot \gamma_2(h_2) \quad \text{și distanța redusă: } h_2 = \sqrt{\left(\frac{h_x}{r_{x_2}}\right)^2 + \left(\frac{h_y}{r_{y_2}}\right)^2} \quad (1.67)$$

- *al treilea model*: tot cu plafon, de tip sferic, cu razele diferite pe cele două direcții, de forma generală:

$$\gamma(h) = w_3 \cdot \gamma_3(h_3) \quad \text{și distanța redusă: } h_3 = \sqrt{\left(\frac{h_x}{r_{x_3}}\right)^2 + \left(\frac{h_y}{r_{y_3}}\right)^2} \quad (1.68)$$

Modelul echivalent complet este obținut prin combinarea celor trei modele elementare și are forma:

$$\gamma(h) = w_1 \cdot \gamma_1(h'_1) + w_2 \cdot \gamma_2(h'_2) + w_3 \cdot \gamma_3(h'_3) \quad (1.69)$$

De reținut că, într-un astfel de model geometric, toate componentele variogramelor direcționale trebuie să aibă același plafon și același model în toate direcțiile.

În practică rar se întâlnește o *anizotropie zonală pură*, adică o rază de influență constantă a variogramei în raport cu direcția la o variație direcțională a plafonului acesteia. Foarte frecventă este compunerea *anizotropiei geometrice cu cea zonală*, ceea ce determină atât variația razei de influență cât și a plafonului variogramei. Exemplul din **Fig.1.60** constă din trei modele de variograme direcționale, fiecare fiind constituit dintr-un singur model elementar. Modelele direcționale de-a lungul axelor Ox și Oy au același plafon și raze diferite (anizotropie geometrică), iar modelul de-a lungul axei Oz are o rază mai scurtă și un plafon mai mare decât modelele direcționale pe Ox și Oy .

Modelul izotrop echivalent este format din două structuri:

- o structură cu anizotropie geometrică formată din primele două modele pe axele Ox și Oy ;
- o structură cu anizotropie zonală formată din modelul pe axa Oz .

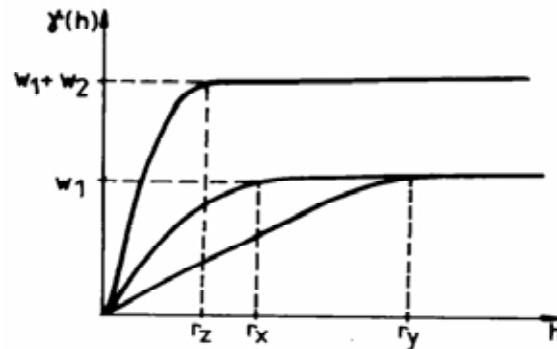


Fig.1.60. Un model de anizotropie zonală și geometrică

Pentru prima structură, modelul echivalent va fi un model izotrop cu plafonul w_1 și raza unitară, ecuația lui echivalentă fiind:

$$\gamma(h) = w_1 \gamma_1(h_1), \text{ cu } \textit{distanța redusă}: h_1 = \sqrt{\left(\frac{h_x}{r_x}\right)^2 + \left(\frac{h_y}{r_y}\right)^2 + \left(\frac{h_z}{r_z}\right)^2} \quad (1.70)$$

Pentru cea de-a doua structură, plafonul este egal cu w_2 și există numai în direcția Oz . Această componentă zonală este modelată cu ecuația:

$$\gamma(h) = w_2 \gamma_2(h_2), \text{ cu } \textit{distanța redusă}: h_2 = \frac{h_z}{r_z} \quad (1.71)$$

Modelul complet este dat de ecuația:

$$\gamma(h) = w_1 \gamma_1(h_1) + w_2 \gamma_2(h_2) \quad (1.72)$$

Analiza variografică se aplică în mod analog variabilelor de tip *numeric* și *alfanumeric*. Diferențieri apar la:

- *codificarea informației primare*, în care *variabilele alfanumerice* sunt codificate doar prin două valori numerice (zero și unu) în timp ce *variabilele numerice* pot lua o infinitate de valori cuprinse într-un interval delimitat de valoarea minimă și maximă;
- *interpretarea legii de variație spațială* exprimată de variogramă; în cazul variabilelor alfanumerice *variograma indicatoare* cuantifică legitatea pe baza căreia se evaluează *probabilitatea* de identificare a variabilei într-un anumit punct iar în cazul variabilelor numerice *variograma experimentală* cuantifică legitatea pe baza căreia într-un anumit punct se evaluează *valoarea* acestora.

Ca diferențiere formală, variograma calculată pentru variabilele alfanumerice poartă denumirea de *variogramă indicatoare* iar pentru cele numerice *variogramă experimentală*.

A14.Variograma indicatoare

Să se evalueze variograma indicatoare a argilei și nisipului de-a lungul unei secțiuni de 40 m, cercetată prin cinci foraje de explorare de 35 m adâncime, amplasate echidistant (Fig.1.61).

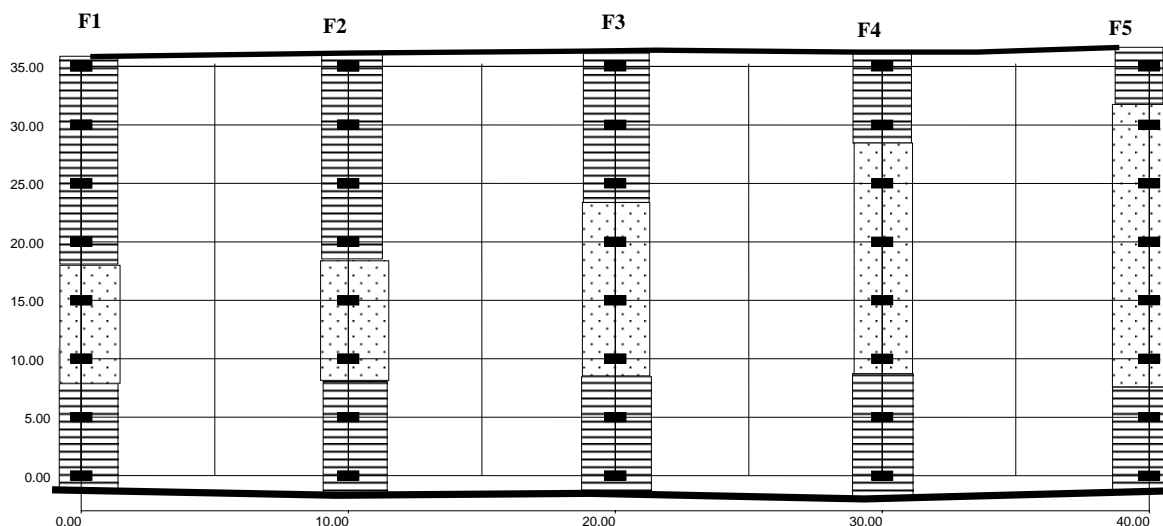


Fig.1.61.Datele primare necesare realizării secțiunii litologice

Rezolvare:

Sucesiunea litologică traversată este constituită numai din argile și nisipuri slab consolidate iar stabilirea variogramelor indicatoare necesită parcurgerea următoarelor etape:

1.Preprocesarea datelor primare

Obiectivul acestei etape este discretizarea și codificarea datelor primare care în cazul studiat sunt constituite din succesiunile litologice continui traversate de cele cinci foraje.

Discretizarea informației presupune transformarea succesiunilor litologice continui din cele cinci foraje într-un șir finit de valori calitative. Pasul/intervalul (e) de discretizare se alege în funcție de complexitatea succesiunii litologice și distanța dintre foraje, iar în cazul studiat se propune a fi $e = 5m$.

Pentru discretizarea datelor primare se amplasează un sistem de referință rectangular în care să fie plasate punctele de discretizare. Originea sistemului de referință are ordonata la cota zero și abscisa pe verticala forajului F1. Datele discretizate sunt sintetizate în tabelul 1.21. Ele sunt reprezentate prin 40 de puncte (5 foraje x 8 puncte pentru fiecare foraj) de coordonate (x,y) cunoscute în care se știe litologia (argilă/nisip).

Codificarea datelor discretizate presupune transformarea valorilor calitative “argilă” și “nisip” în valori numerice adaptate tehnicilor de prelucrare cantitativă. Valorile numerice utilizate pentru codificarea succesiunii litologice discretizate din tabelul 1.21 sunt:

- 0 (zero) pentru absența tipului litologic studiat pe poziția punctului din rețeaua de discretizare a datelor primare;
- 1 (unu) pentru prezența tipului litologic studiat.

Tabelul 1.21. Rezultatele discretizării datelor primare

Forajul F1			Forajul F2			Forajul F3			Forajul F4			Forajul F5		
x	y	Litologie	x	y	Litologie	x	y	Litologie	x	y	Litologie	x	y	Litologie
0	0	argila	10	0	argila	20	0	Argila	30	0	argila	40	0	argila
0	5	argila	10	5	argila	20	5	Argila	30	5	argila	40	5	argila
0	10	nisip	10	10	nisip	20	10	Nisip	30	10	nisip	40	10	nisip
0	15	nisip	10	15	nisip	20	15	Nisip	30	15	nisip	40	15	nisip
0	20	argila	10	20	argila	20	20	Nisip	30	20	nisip	40	20	nisip
0	25	argila	10	25	argila	20	25	Argila	30	25	nisip	40	25	nisip
0	30	argila	10	30	argila	20	30	Argila	30	30	argila	40	30	nisip
0	35	argila	10	35	argila	20	35	Argila	30	35	argila	40	35	argila

Procedând după aceste reguli de discretizare, cu datele din **tabelul 1.21** se obțin datele codificate pentru calculul *variogramei indicatoare* (**tabelul 1.22**). Pentru prelucrarea automată a acestora cu pachetul de programe Geo-EAS este necesară organizarea lor într-un fișier de tip ASCII cu structură specială (*Cap.4*).

Tabelul 1.22. Datele codificate necesare calculului *variogramei indicatoare*

Nr.	x	y	Argila	Nisip	Litologie	Nr.	x	Y	Argila	Nisip	Litologie
1	0	0	1	0	argila	21	20	20	0	1	nisip
2	0	5	1	0	argila	22	20	25	1	0	argila
3	0	10	0	1	nisip	23	20	30	1	0	argila
4	0	15	0	1	nisip	24	20	35	1	0	argila
5	0	20	1	0	argila	25	30	0	1	0	argila
6	0	25	1	0	argila	26	30	5	1	0	argila
7	0	30	1	0	argila	27	30	10	0	1	nisip
8	0	35	1	0	argila	28	30	15	0	1	nisip
9	10	0	1	0	argila	29	30	20	0	1	nisip
10	10	5	1	0	argila	30	30	25	0	1	nisip
11	10	10	0	1	nisip	31	30	30	1	0	argila
12	10	15	0	1	nisip	32	30	35	1	0	argila
13	10	20	1	0	argila	33	40	0	1	0	argila
14	10	25	1	0	argila	34	40	5	1	0	argila
15	10	30	1	0	argila	35	40	10	0	1	nisip
16	10	35	1	0	argila	36	40	15	0	1	nisip
17	20	0	1	0	argila	37	40	20	0	1	nisip
18	20	5	1	0	argila	38	40	25	0	1	nisip
19	20	10	0	1	nisip	39	40	30	0	1	nisip
20	20	15	0	1	nisip	40	40	35	1	0	argila

Numărul de variabile prelucrabile din fișier este 5 (*numărul curent al punctului, x, y, Argilă, Nisip*). Ultima coloană, cu *Litologia*, este un șir de *valori alfanumerice (argilă/nisip)* pe care programul Geo-EAS nu le prelucrează, sunt plasate în fișier pe coloana a șasea și sunt utile pentru verificarea codificării.

2.Evaluarea ponderii valorilor variabilei alfanumerice

Pentru evaluarea ponderii valorilor variabilelor alfanumerice (*litologia*) se utilizează *histograma nominală* (**Fig.1.62**) care conține pe axa orizontală *numele* variabilei iar pe verticală *frecvența absolută* a acestora.

Histograma nominală a litologiei evidențiază predominanța argilei (24 valori) față de nisip (16 valori).

Cu ajutorul histogramei nominale se stabilește ordinea în care se realizează estimarea distribuției spațiale a valorilor variabilei, în special în cazul unui număr mai mare de valori alfanumerice.

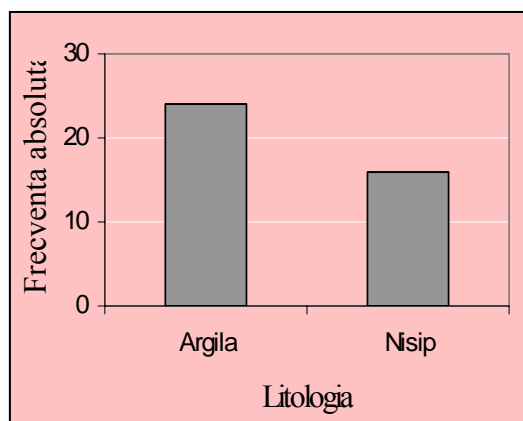


Fig.1.62. Histograma nominală.

3. Identificarea unităților structurale

Identificarea unităților structurale se face prin analiza frecvenței de apariție a fiecărui litotip (argilă/nisip) atât pe verticală cât și pe orizontală. În acest scop, pentru fiecare abscisă și ordonată a grilei de discretizare a datelor primare se stabilesc frecvențele absolute de apariție ale argilei și nisipului (**tabelul 1.23** și **1.24**).

Tabelul 1.23. Frecvențele de apariție pe verticală ale argilei și nisipului

y [m]	Frecvența absolută	
	Argilă	Nisip
0	5	0
5	5	0
10	0	5
15	0	5
20	2	3
25	3	2
30	4	1
35	5	0

Tabelul 1.24. Frecvențele de apariție pe orizontală ale argilei și nisipului

x [m]	Frecvența absolută	
	Argilă	Nisip
0	6	2
10	6	2
20	5	3
30	4	4
40	3	5

Reprezentarea grafică a acestor frecvențe permite precizarea următoarelor observații:

- pe verticală.:
 - *argila* este distribuită în două unități structurale (**Fig.1.63a**):
 - unitatea inferioară, cu grosime constantă dezvoltată între +0 și +10 m;
 - unitatea superioară, cu grosime variabilă prezentă între +20 și +40 m.
 - *nisipul* este distribuit într-o singură unitate structurală cuprinsă între +15 și +35m (**Fig.1.63b**);

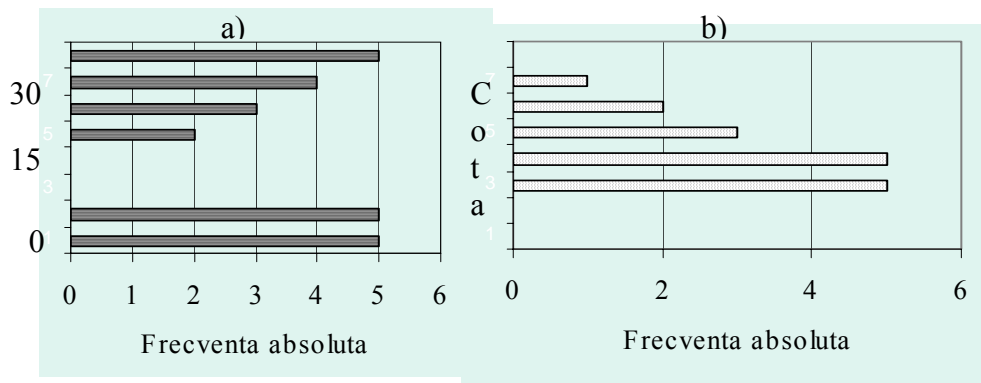


Fig. 1.63. Frecvențele de distribuție ale argilei și nisipului pe verticală

- pe orizontală:
 - argila, atât în unitatea inferioară cât și în cea superioară, are o variație lentă care nu justifică separarea unor alte unități structurale pe orizontală (Fig. 1.64a);
 - nisipul are aceeași comportare ca argila (Fig. 1.64b)

În concluzie, analiza distribuției spațiale atât pe orizontală cât și pe verticală a celor doi litotipi a identificat trei unități structurale:

- două pentru argilă;
- una pentru nisip.

Fiecare din aceste structuri se tratează separat în etapa de calcul și modelare a variogramelor indicatoare.

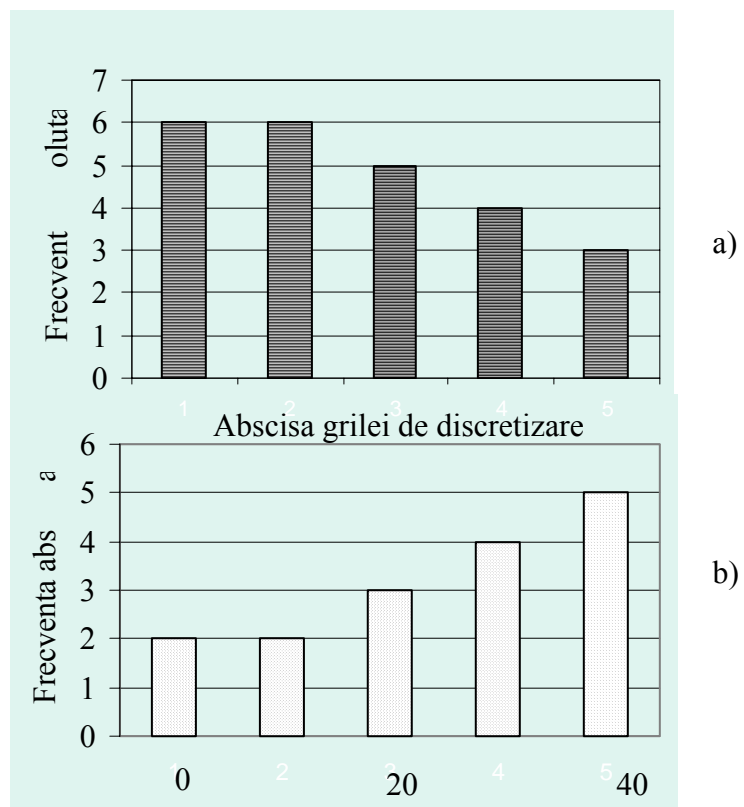
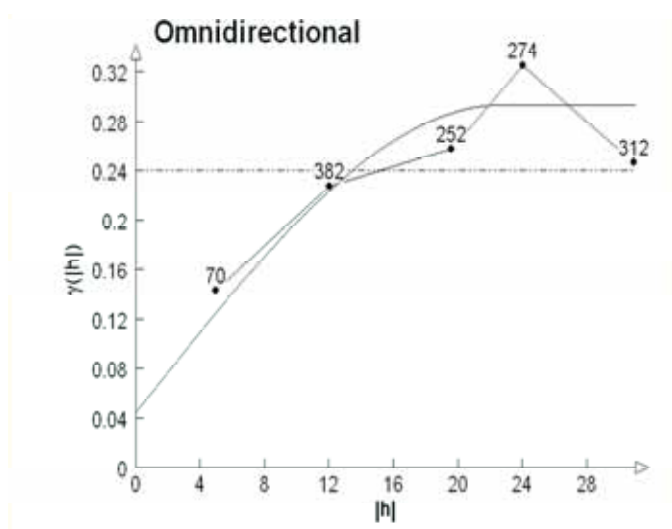


Fig. 1.64. Frecvențele de distribuție ale argilei și nisipului pe orizontală

4. Calculul și modelarea variogramelor indicatoare

Calculul și modelarea variogramelor indicatoare se face separat pentru unitățile litologice și structurale identificate. Particularitatea unității structurale inferioare a argilei (grosimea constantă a acesteia) face necesar calculul variogramei indicatoare doar pentru unitatea structurală superioară a argilei și cea a nisipului.



Unitatea structurală superioară a argilei are o variogramă indicatoare izotropă (Fig.1.65) de tip sferic cu parametrii:

- efectul de pepită=0,045;
- palierul = 0,25;
- raza de influență =23 m.

Fig.1.65. Modelul variogramei omnidirecționale pentru unitatea structurală superioară a argilei.

Unitatea structurală a nisipului se manifestă de asemenea izotrop, variograma indicatoare fiind tot de tip sferic cu parametrii:

- efectul de pepită = 0,08;
- palierul = 0,2;
- raza de influență = 20 m.

O analiză comparativă a celor două modele de variogramă pentru cei doi litotipi permite formularea următoarelor observații:

- *gradul de continuitate* al celor doi litotipi este identic, reflectat de același model de variogramă, fapt care poate fi interpretat ca o menținere a condițiilor de sedimentare pe parcursul depunerii întregii secvențe litologice studiate;
- *mărimea erorii* cu care a fost stabilită legitatea variabilității spațiale a celor doi litotipi, cuantificată prin valoarea palierului celor două modele (0,25 pentru argilă, unitatea structurală superioară și 0,2 pentru nisip), este aceeași, lucru care indică un grad de complexitate analog pentru cei doi litotipi;
- *amplimea dezvoltării spațiale* a celor două unități structurale, reflectată de raza de influență (23 m pentru argilă 20 m pentru nisip), este similară, fapt ce confirmă corectitudinea utilizării aceleiași densități a punctelor de observație pentru cercetarea distribuției ambilor litotipi.

Din punct de vedere practic, toate aceste observații conduc la concluzia că în etapa de construire a secțiunii litologice, fără a se comite erori semnificative, se poate utiliza același model de variogramă atât pentru unitatea structurală superioară a argilei cât și pentru unitatea structurală a nisipului, adică: *tip sferic cu efectul de pepită 0,06, palierul 0,22 și raza de influență 22 m.*

COMENTARIU

Analiza variografică a variabilelor calitative este un moment de revelație pentru mulți din practicienii Geostatisticii. Modul de formulare a problemei și interpretare a rezultatelor te scoate din rutina prelucrărilor numerice conexe estimărilor spațiale.

Pentru cercetarea extinderii unui orizont de argile roșii într-un areal oarecare formularea problemei ar fi cam așa:

Dacă în 12 din 25 de aflorimente cartate și în 11 din 50 de foraje realizate găsim argila roșie, care este șansa să găsim argila roșie în al 26-lea afloriment sau în al 51-lea foraj pe care urmează să le cercetăm?

Unde trebuie să caut aflorimentul sau să execut forajul pentru a crește șansa de a identifica argila roșie??

*Aceasta este problema centrală a cercetării geologice și cu asta s-au ocupat generații de geologi până au identificat resursele minerale exploatare în prezent. În fazele de pionierat, prelucrarea și interpretarea tuturor informațiilor necesare pentru a soluționa problema se bazau în mare parte pe **analogii** și **intuiția** cercetătorului. Tehnica soluționării problemei era “interzisă” începătorilor care priveau cu uimire și respect cum “bătrânii” care o posedau mergeau pe teren zile întregi îngândurați și tăcuți pentru a pune dintr-o dată, pe o hartă, rezultatul spectaculos al “prelucrărilor”. Așa au fost clarificate multe din misterele structurilor geologice.*

Analiza variografică a variabilelor calitative descifrează o parte din acest mecanism “interzis” și complex care permite evaluarea probabilității ca o anumită formațiune să poată fi identificată într-un anumit punct din spațiu. Ea stabilește pe baza unor observații calitative (ex.: “Am găsit argila roșie lângă fântâna lui Papură !”) o legitate cantitativă de distribuție a unei caracteristici pe baza căreia va fi evaluată șansa de a identifica într-un anumit punct variabila căutată (ex.: “N-ai nici o șansă să găsești argilă la salcâmul lui Nil!”).

Interpretarea acestor rezultate cantitative se face în termeni probabilistici. Rezultatele cantitative ale analizei variografice se interpretează pe două niveluri:

- *cel global (pe baza histogramei nominale; **Fig.1.62**), de unde rezultă dacă șansele de a găsi sunt mai mari decât cele de a nu găsi ce căutăm (ex.: din **Fig.1.62** rezultă că avem șanse mai mari să găsim argilă, decât să nu găsim);*
- *cel spațial (pe baza histogramei orizontale - **Fig.1.63** și verticale - **Fig.1.64**), de unde rezultă șansa de a identifica ce căutăm într-un anumit punct din spațiu (ex.: din **Fig.1.63a** rezultă că șansele de a identifica argila roșie sunt mai mari în partea superioară a secțiunii și mai precis spre vestul acesteia - **Fig.1.64a**).*

Analiza variografică a variabilelor calitative este un instrument inutil în mâna unui cercetător care este străin de modelele conceptuale ale proceselor cercetate. Delimitarea domeniilor de extindere a zonelor cercetate, separarea acestora în subzone omogene, corelarea scării la care se fac determinările cu obiectivul cercetării, toate acestea asigură cadrul aplicării corecte a analizei variografice pentru variabilele calitative.

A15. Variograma omnidirecțională

Să se evalueze variograma omnidirecțională a temperaturilor maxime din august 1998 (tabelul 1.25) înregistrate în 171 de stații meteorologice de pe teritoriul României.

Tabelul 1.25 Temperaturi maxime [°C] în luna august 1998 (România).

x	y	tmax	x	y	tmax	x	y	tmax	x	y	tmax	x	y	tmax
16.1	3.4	29	15.9	1.1	31	19.2	7.9	28	16.6	11.0	23	15.0	6.9	18
16.4	3.0	30	14.4	0.4	32	18.2	7.5	21	14.3	11.5	15	14.6	7.1	8
23.5	6.1	27	14.6	1.4	32	16.4	7.9	28	14.8	12.0	21	14.1	6.9	18
24.5	6.1	26	13.0	0.7	32	17.9	18.7	24	12.4	11.2	23	13.5	6.1	24
25.8	6.1	26	13.3	2.1	31	19.9	8.6	27	20.6	11.5	25	15.3	5.7	25
24.6	6.9	27	13.9	3.3	29	17.1	8.4	13	19.2	11.6	23	14.6	4.9	27
25.7	5.3	26	11.6	2.1	31	15.6	8.9	23	19.9	12.2	25	14.9	3.9	29
22.6	4.6	26	10.6	0.7	32	15.0	7.9	22	10.6	12.5	24	15.0	2.5	29
21.3	4.3	27	7.8	1.6	32	15.0	9.4	23	17.0	12.5	23	6.2	12.0	23
22.2	2.5	27	8.9	1.7	31	16.5	9.1	23	16.0	12.7	13	6.9	11.1	22
23.3	2.3	24	10.3	0.7	30	19.2	9.5	27	16.0	13.1	22	6.3	6.5	19
22.7	2.2	28	7.8	3.4	31	21.0	9.3	25	14.5	12.6	21	7.1	6.7	17
22.9	1.0	26	7.2	4.2	31	20.4	9.9	26	17.1	13.6	23	9.9	7.0	19
21.5	1.8	27	8.5	3.7	30	17.7	10.0	26	13.8	13.9	14	9.1	6.2	15
21.3	6.3	27	11.7	3.3	31	15.6	10.9	21	20.2	13.5	25	9.1	7.0	25
19.5	5.7	28	11.4	4.1	29	14.3	10.2	22	18.6	13.5	24	7.9	7.5	23
19.3	4.5	27	10.6	4.0	29	12.8	7.7	13	18.4	14.5	23	5.3	8.1	25
19.8	3.8	29	13.1	4.7	27	12.5	7.0	20	17.6	14.5	23	3.6	8.5	26
20.8	3.0	27	12.2	5.4	24	11.8	6.7	23	14.9	14.8	24	11.5	11.8	22
20.0	3.1	28	11.6	5.8	28	11.7	6.4	16	15.5	16.0	22	6.0	11.6	21
19.8	2.3	29	10.0	4.8	29	13.5	8.5	23	16.8	15.3	23	9.0	10.4	16
17.8	1.9	30	10.4	6.0	24	12.5	9.1	22	17.1	16.4	24	8.4	10.6	22
18.8	3.2	29	8.8	5.6	28	11.2	8.4	22	17.8	15.6	23	10.4	10.3	25
17.7	3.0	29	7.8	5.5	30	9.8	9.0	22	18.6	16.8	22	10.8	9.8	24
17.8	4.2	28	6.5	5.2	28	8.0	8.6	24	18.9	16.2	24	9.6	9.5	24
18.3	5.9	28	4.5	4.6	25	7.6	9.5	23	10.4	11.3	21	12.6	9.9	23
18.8	6.2	28	5.4	5.4	26	5.8	9.4	25	9.1	11.3	15	19.7	10.9	25
19.8	7.2	27	4.6	5.7	25	2.0	9.8	27	9.8	12.1	21	18.5	11.2	24
21.4	7.1	27	5.5	6.1	18	3.8	10.0	25	11.0	13.6	23	8.4	12.6	21
17.0	6.4	25	5.2	6.1	25	4.6	10.2	22	12.3	13.4	22	7.5	12.1	13
16.1	5.9	28	4.5	7.2	25	5.2	10.9	22	12.6	15.1	10	5.5	13.3	24
16.0	6.9	19	3.3	7.1	27	4.3	11.5	24	10.7	16.4	22	4.8	12.4	23
14.9	6.5	16	6.1	7.1	25	6.5	10.3	24	10.7	16.0	21	7.4	13.1	23
6.1	14.4	22	8.5	13.7	22	7.8	14.8	23	10.0	15.5	22	7.3	15.7	23
												8.2	16.0	23

Rezolvare:

Este *numărul* esența tuturor lucrurilor? Poate *numărul* să exprime totul?

Privind **tabelul 1.25**, plin cu numere, vom da un răspuns negativ la ambele întrebări.. Ce ne pot spune aceste numere despre distribuția spațială a temperaturilor maxime din 28 august 1998?

Formatul tabelar al datelor este total inexpressiv și nu ne ajută să ne formăm o

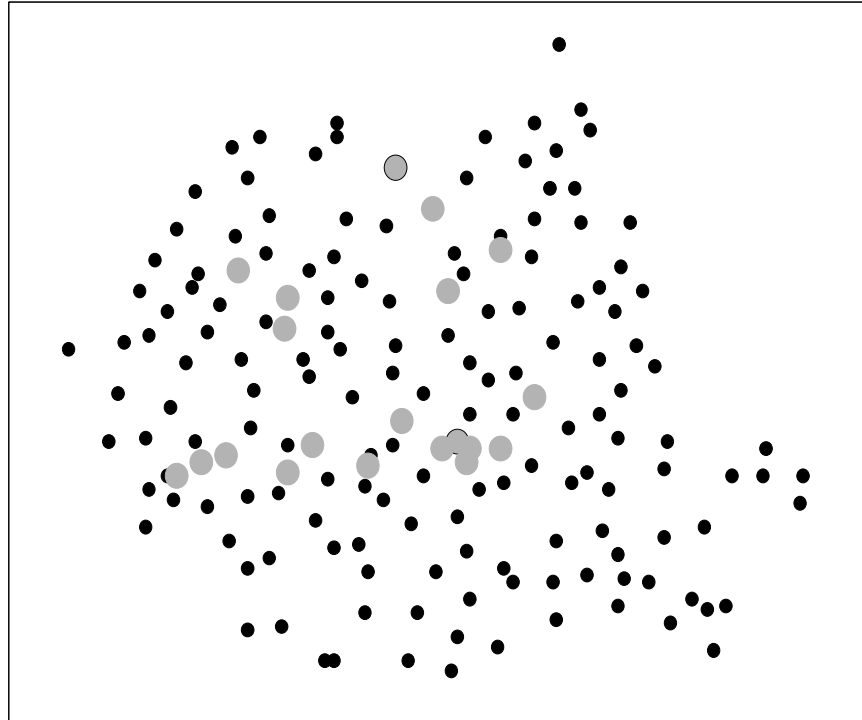


Fig.1.66. Distribuția stațiilor meteorologice

imagine sugestivă în legătură cu subiectul întrebării. Prelucrate după algoritmi adecvați *numerele* își dezvăluie *puterea*, *absolutizată* de unii sau *ignorată* pe nedrept de alții.

Prin calculul *variogramei omnidirecționale* încercăm acum să extragem din *numerele* sintetizate în **tabelul 1.25** informații despre distribuția spațială a temperaturii sau, altfel spus:

Cum variază temperatura de la o stație meteorologică la alta în 28 august 1998, în România?

Pentru estimarea corectă a *variogramei omnidirecționale* trebuie analizate:

- distribuția spațială a stațiilor în care au fost măsurate temperaturile;
- variabilitatea globală a temperaturilor.

Distribuția stațiilor meteorologice are o densitate relativ *constantă* pe întreg teritoriul României (**Fig.1.66**) iar numărul mare de stații permite calculul *variogramei omnidirecționale* cu o precizie satisfăcătoare.

Analiza variabilității globale a temperaturii din cele 171 de stații meteorologice indică o selecție de date *neomogenă* cu *repartiție bimodală* (**Fig.1.67**) care implică separarea selecției de valori în două grupe:

- *grupa 1*, formată din 151 de valori măsurate în stațiile meteorologice amplasate în zonele de câmpie și deal (discuri negre în **Fig.1.66**);
- *grupa 2*, formată din 20 de valori măsurate în stații meteorologice situate în zonele montane (discuri gri în **Fig.1.66**).

Pentru calculul *variogramei omnidirecționale* se utilizează în cadrul acestei aplicații cele 151 de valori din *grupa 1*, valori care dau structura spațială generală a temperaturii aerului din România la momentul efectuării măsurătorilor.

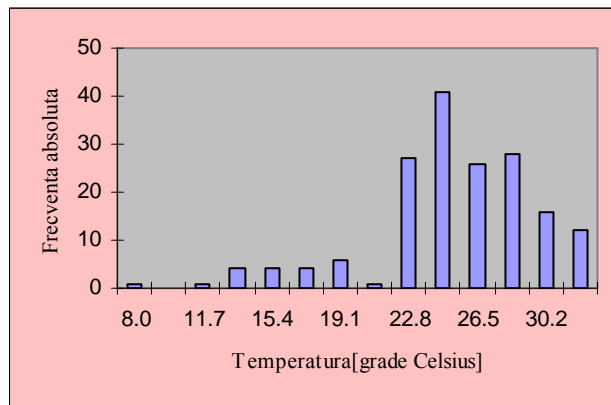


Fig.1.67 Histograma celor 171 de valori ale temperaturii.

Tabelul 1.26 Valorile variogramei omnidirecționale

Nr	N(h)	h_{med}	$\gamma(h)$
0	60	0.70	1.47
1	1202	1.82	2.14
2	2076	3.47	3.96
3	2660	5.13	6.06
4	3006	6.81	8.28
5	3080	8.51	10.29
6	2888	10.20	11.85
7	2542	11.87	12.83
8	2080	13.57	13.33
9	1496	15.25	13.17

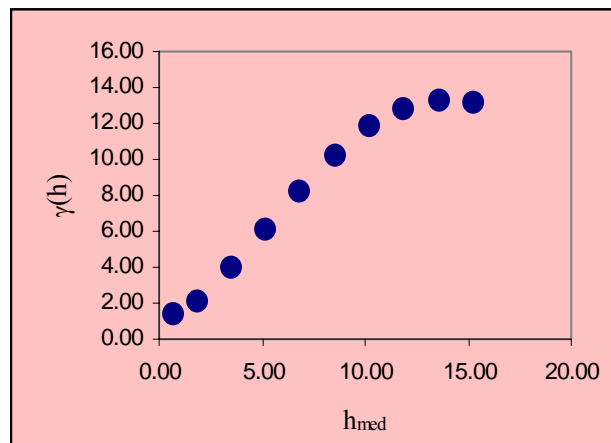


Fig.1.68. Variograma omnidirecțională a temperaturii.

Variograma omnidirecțională este calculată pentru 10 clase de distanțe (**tabelul 1.26**) și indică o foarte bună continuitate a temperaturii (**Fig.1.68**) care poate fi modelată cu un model gaussian (**Fig.1.69**) cu parametrii:

- efect de pepită = 1,24;
- palierul = 12,35
- raza de influență = 12,8

Utilizarea *variogamei omnidirecționale* ca lege de variabilitate spațială a temperaturii presupune acceptarea ipotezei unei structuri izotrope pentru care *variograma de suprafață* este

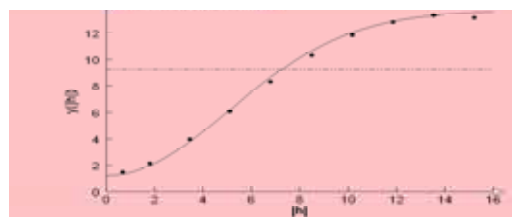


Fig.1.69. Modelul variogramei

constituită din cercuri concentrice (**Fig.1.70**).

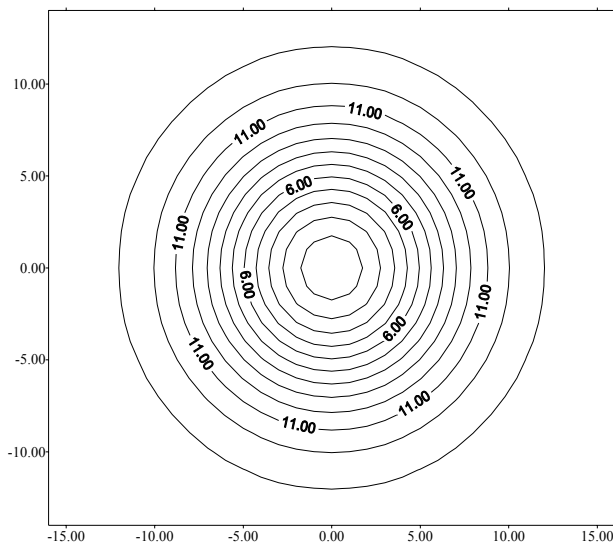


Fig.1.70. Variograma de suprafață a temperaturii în cazul unei variograme omnidirecționale

Existența unei variograme omnidirecționale de tip gaussian este premiza favorabilă realizării unei hărți a distribuției temperaturii cu o precizie satisfăcătoare pe toate direcțiile de interpolare.

Modelul structural identificat este utilizat în etapa de estimare a distribuției spațiale.

“Absența” variogramei omnidirecționale indică imposibilitatea identificării unei legi de distribuție pe baza datelor disponibile. “Absența” variogramei omnidirecționale constă în lipsa corelației dintre valorile variogramei ($\gamma(h)$) și distanțele pentru care acestea sunt calculate (h).

COMENTARIU

Obținerea unei variograme omnidirecționale diferită de cea denumită efect de pepiță total este tot ce-și dorește un geostatistician pentru că:

Dacă variogramă omnidirecțională nu e, nimic nu e !

Dacă variogramă omnidirecțională nu e geostatisticianul intră în panică pentru că din acel moment se simte inutil. Când a ajuns la această concluzie își poate lua jucăriile și poate să plece acasă. Ca să nu rămână fără “job” face toate eforturile, iar direcțiile în care poate acționa sunt:

- *completarea selecției de date cu valori suplimentare pe baza cărora să se poată identifica legea de variație spațială a variabilei studiate (variograma omnidirecțională);*
- *testarea influenței fiecărei valori sau grup de valori asupra funcției de variogramă (pentru eliminarea valorilor aberante care pot masca legea căutată a variației spațiale);*
- *renunțarea la evaluarea distribuției spațiale a variabilei studiate și calculul unei valori medii pentru întreg arealul probat, prin metodele analizei variabilității globale.*

Nici o variantă de lucru nu trebuie neglijată în această etapă de prelucrare care are ca obiectiv optimizarea utilizării informației disponibile.

A16. Analiza anizotropiei structurilor spațiale

Este structura spațială a temperaturilor maxime din 28 august 1998 pe teritoriul României izotropă? Legea de variabilitate spațială a acestora este aceeași în toate direcțiile? Factorii generali (latitudinea) și particulari (relieful) își fac simțită influența asupra legii de distribuție spațială a temperaturii?

Rezolvare:

Pentru a răspunde la aceste întrebări apelăm la:

- variograma de suprafață;
- variogramele direcționale.

Variograma de suprafață a celor 151 de temperaturi din grupa 1, separată în aplicația A15, indică o pronunțată anizotropie cu direcția de variabilitate minimă orientată V-E (Fig. 1.71). Este ușor de intuit că această anizotropie este determinată de poziția stațiilor meteorologice în raport cu ecuatorul (latitudinea) iar neregularitățile care perturbă liniile conturale sunt determinate în principal de relief.

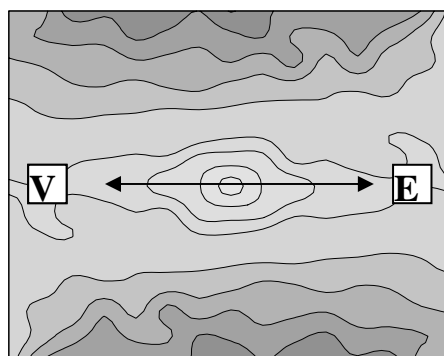


Fig. 1.71. Variograma de suprafață a temperaturilor maxime din 28 august 1998.

Este momentul să ne reamintim că variograma de suprafață este utilă și în analiza staționarității variabilelor regionalizate (vezi aplicația A12). Morfologia variogramei de suprafață sugerează prezența unei tendințe regionale care se manifestă pe direcția N-S. Prezența acestei tendințe poate fi confirmată prin calculul variogramelor direcționale iar legitatea acesteia poate fi studiată cu suprafețele polinomiale de tendință (vezi aplicația A13).

Direcțiile de calcul pentru variogramele direcționale se stabilesc din morfologia variogramei de suprafață, ele fiind:

- direcția de variabilitate minimă, VE, în cazul temperaturilor maxime din 28 august 1998;
- direcția de variabilitate maximă, NS, în cazul temperaturilor maxime din 28 august 1998.

Variograma direcțională pe direcția de variabilitate minimă VE (Fig. 1.72) indică o continuitate medie, de model sferic, cu tendința de stabilizare la valori ale palierului de 4,7, corespunzător unei raze de influență cuprinsă între 10 și 12 unități grafice.

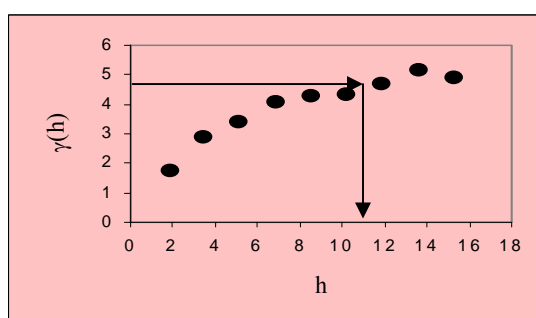


Fig. 1.72. Variograma direcțională pe direcția VE (toleranță unghiulară=20 grade sexagesimale).

Direcția VE este direcția pe care se stabilește modelul de variogramă al structurii câmpului temperaturilor. Pe această direcție se evaluează raza maximă de influență a modelului structurii spațiale, rază pe baza căreia se stabilește distanța optimă dintre punctele de observație ale unei rețele de monitoring.

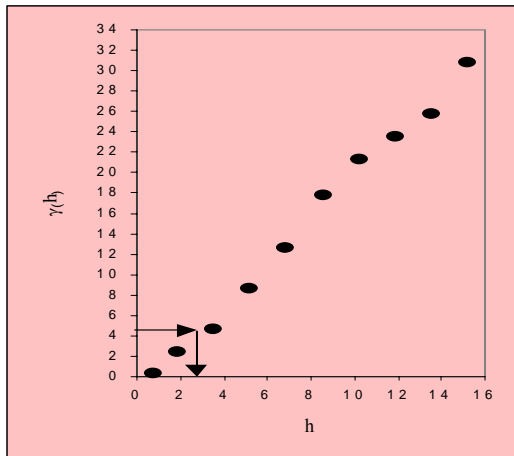


Fig.1.73. Variograma direcțională pe direcția NS (toleranța unghiulară de 20 grade sexagesimale).

sferic, având:

- efectul de pepită = 0;
- palierul = 4,7;
- razele de influență: $r_{V-E} = 10 - 12$; $r_{N-S} = 2 - 4$

Pe direcția NS, variograma direcțională (Fig.1.73) indică o evidentă nestaționaritate a câmpului temperaturilor prin creșterea continuă a valorilor variogramei proporțional cu creșterea distanței de calcul (h).

Pentru palierul de 4,7 stabilit pentru variograma pe direcția VE, raza de influență corespunzătoare direcției NS este cuprinsă între 2 și 4 unități grafice. În aceste condiții raportul de anizotropie al structurii temperaturii este cuprins între 3 și 5 ($12/4$ și $10/2$).

Analiza structurii spațiale a temperaturilor din 28 august 1998 conduce la concluzia că anizotropia acesteia este de tip geometric, cu model

COMENTARIU

Trebuie reținut că analiza anizotropiei structurilor se declanșează numai după ce s-a obținut o variogramă omnidirecțională (calculată cu o toleranță de 90 grade sexagesimale) care indică prezența unei structuri spațiale.

Prezența unei structuri spațiale este semnalată de obținerea unor variograme omnidirecționale de model gaussian, sferic, exponențial sau linear. Obținerea unei variograme efect de pepită total semnalează lipsa unei structuri spațiale și determină necesitatea completării informației sau pe aceea a realizării unor prelucrări suplimentare.

Niciodată nu se începe analiza anizotropiei după obținerea unei variograme efect de pepită total!!!

Niciodată nu se renunță la prelucrare după primul insucces... adică după obținerea primei variograme omnidirecționale efect de pepită total! Dar aceasta nu înseamnă că se continuă la infinit prelucrarea sterilă o unor date care se dovedesc nereprezentative pentru structura studiată. Trebuie să ne formăm un "bun simț al insistenței" la care se ajunge prin exercițiul prelucrării multor categorii de date.

Trebuie să extragem informația din datele existente fără să le denaturăm prin filtrări "abuzive" sau "tendențioase". Este periculos să "aranjăm" datele ca să obținem ce ne dorim! Aplicate corect, criteriile statistice proprii metodelor probabiliste ne permit respectarea semnificației reale a datelor primare.

Analiza variografică este etapa în care se poate denatura cu mare ușurință semnificația datelor primare. Aceasta este etapa care pune la încercare profesionalismul geostatisticianului.

A17. Validarea modelului de variogramă

Să se valideze modelul de variogramă al grosimii unui strat de nisip explorat prin 121 de foraje amplasate într-o rețea pătratică cu parametrul 100 m (Fig.1.74, tabelul 1.27).

Rezolvare:

Validarea modelului de variogramă este o verificare a acestuia pe baza valorilor măsurate în punctele de observație disponibile. Ideea care stă la baza validării este următoarea: modelul de variogramă este “bun” dacă utilizându-l putem obține prin calcul valorile variabilei măsurate în punctele de observație.

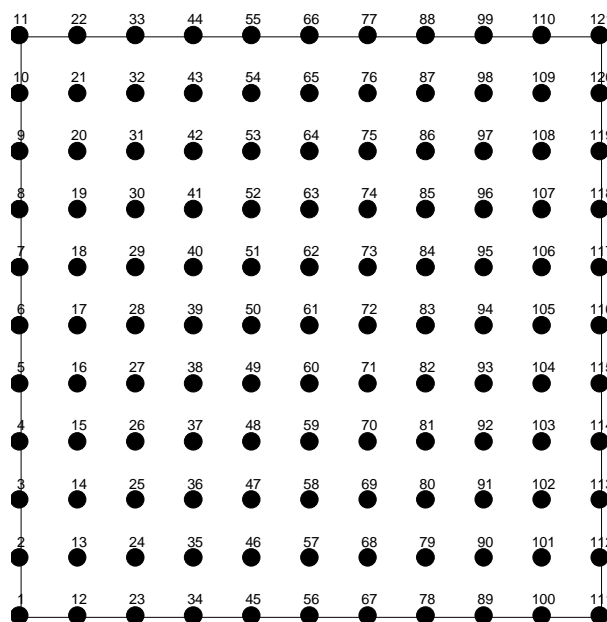


Fig.1.74. Distribuția forajelor de explorare

Este ușor de intuit că dacă modelul este perfect, prin calcul vom putea obține exact valorile care au fost măsurate în punctele de observație. Dar cum nimic nu este perfect, evident că valorile obținute prin calcul cu ajutorul modelului vor diferi față de cele măsurate. Optimizarea modelului de variogramă tinde să minimizeze diferența dintre valoarea calculată cu ajutorul modelului (v_i^*) și cea măsurată în punctul de observație (v_i).

În plus, trebuie reținut că și un model perfect nu ne asigură perfecțiunea estimărilor în punctele din vecinătatea punctelor de observație... sau altfel spus: *nici o metodă de interpolare nu este perfectă* deoarece *nici o probare nu poate fi exhaustivă* (...și dacă ar fi probarea exhaustivă, estimarea nu mai are justificare).

Modelul de variogramă pe care ne propunem să-l validăm este unul de anizotropie geometrică având următorii parametri:

- model sferic;
- efect de pepită = 0;
- palier = 0,075;
- $r_1 = 500$ m, pe direcția de continuitate maximă cu orientarea: $\theta = 45^\circ$;
- $r_2 = 300$ m, pe direcția de continuitate minimă cu orientarea: $\theta = 135^\circ$.

Deoarece repartiția valorilor grosimii este lognormală calculul variogramei experimentale și modelarea ei s-au realizat cu valori logaritmice care asigură o normalizare aproape perfectă a valorilor grosimii (o reducere a asimetriei de la 1,13 la 0,19).

Pentru calculul valorilor grosimii pe poziția punctelor de observație se utilizează ca metodă de interpolare kriging-ul punctual ordinar deoarece pe distanțe de maximum 500 m variația grosimii este practic staționară. Rezultatele obținute pe baza modelului stabilit pentru structura spațială a stratului de nisip sunt sintetizate în tabelul 1.28 (G - grosimea măsurată; G^* - grosimea calculată).

Tabelul 1.27. Valorile măsurate ale grosimii stratului de nisip.

Nr.	x	y	G	Nr.	x	y	G	Nr.	x	y	G
1	0	0	6.3	41	300	700	5	81	700	300	5
2	0	100	5.5	42	300	800	3	82	700	400	4.5
3	0	200	5	43	300	900	9	83	700	500	4
4	0	300	4.25	44	300	1000	3.5	84	700	600	3.5
5	0	400	3.6	45	400	0	6.3	85	700	700	3.1
6	0	500	3	46	400	100	5.6	86	700	800	3.5
7	0	600	3.4	47	400	200	5	87	700	900	3.5
8	0	700	3.4	48	400	300	4.5	88	700	1000	3.4
9	0	800	3	49	400	400	3.9	89	800	0	6.8
10	0	900	2.5	50	400	500	3.3	90	800	100	6.3
11	0	1000	2	51	400	600	3.5	91	800	200	5.7
12	100	0	6.2	52	400	700	10	92	800	300	5.3
13	100	100	5.4	53	400	800	9	93	800	400	4.6
14	100	200	4.7	54	400	900	5	94	800	500	4.4
15	100	300	4.2	55	400	1000	10	95	800	600	4
16	100	400	3.5	56	500	0	6.4	96	800	700	3.5
17	100	500	3	57	500	100	5.7	97	800	800	3.2
18	100	600	4	58	500	200	5.2	98	800	900	3.2
19	100	700	4.5	59	500	300	4.6	99	800	1000	3.3
20	100	800	4	60	500	400	4	100	900	0	7.2
21	100	900	2.75	61	500	500	3.5	101	900	100	6.5
22	100	1000	2.4	62	500	600	3.4	102	900	200	6
23	200	0	6	63	500	700	4.9	103	900	300	5.6
24	200	100	5.3	64	500	800	2	104	900	400	5.2
25	200	200	4.8	65	500	900	4.5	105	900	500	4.7
26	200	300	4.25	66	500	1000	3.6	106	900	600	4.4
27	200	400	3.6	67	600	0	6.5	107	900	700	4
28	200	500	3.1	68	600	100	5.9	108	900	800	3.5
29	200	600	4	69	600	200	5.3	109	900	900	3.2
30	200	700	7	70	600	300	4.8	110	900	1000	3.2
31	200	800	5	71	600	400	4.2	111	1000	0	7.3
32	200	900	3.5	72	600	500	3.7	112	1000	100	6.8
33	200	1000	2.8	73	600	600	3.1	113	1000	200	6.2
34	300	0	6.2	74	600	700	4	114	1000	300	5.8
35	300	100	5.5	75	600	800	8	115	1000	400	5.4
36	300	200	4.9	76	600	900	3.8	116	1000	500	5
37	300	300	4.4	77	600	1000	3.5	117	1000	600	4.5
38	300	400	3.7	78	700	0	6.6	118	1000	700	4.2
39	300	500	3.25	79	700	100	6.1	119	1000	800	3.8
40	300	600	3.8	80	700	200	5.5	120	1000	900	3.4
								121	1000	1000	3

Tabelul 1.28. Rezultatele validării modelului de variogramă

x	y	G	G*	G-G*	x	y	G	G*	G-G*	x	y	G	G*	G-G*
600	800	2.08	1.22	-0.86	0	100	1.70	1.70	-0.01	600	400	1.44	1.45	0.01
400	1000	2.30	1.45	-0.85	400	400	1.36	1.35	-0.01	500	100	1.74	1.75	0.01
400	800	2.20	1.39	-0.81	200	300	1.45	1.44	-0.01	500	400	1.39	1.40	0.01
400	700	2.30	1.55	-0.75	800	700	1.25	1.25	-0.01	1000	200	1.82	1.84	0.01
300	900	2.20	1.45	-0.75	800	100	1.84	1.84	-0.01	0	700	1.22	1.24	0.02
200	700	1.95	1.47	-0.48	100	300	1.44	1.43	-0.01	700	500	1.39	1.40	0.02
500	900	1.50	1.23	-0.28	200	0	1.79	1.79	0.00	800	200	1.74	1.76	0.02
500	700	1.59	1.44	-0.15	300	100	1.70	1.70	0.00	300	400	1.31	1.33	0.02
100	800	1.39	1.27	-0.12	700	400	1.50	1.50	0.00	1000	600	1.50	1.53	0.02
0	0	1.84	1.77	-0.07	900	800	1.25	1.25	0.00	700	1000	1.22	1.25	0.02
0	200	1.61	1.55	-0.06	700	100	1.81	1.80	0.00	100	400	1.25	1.28	0.03
0	400	1.28	1.23	-0.05	800	0	1.92	1.91	0.00	200	100	1.67	1.69	0.03
1000	0	1.99	1.94	-0.05	1000	700	1.44	1.43	0.00	800	900	1.16	1.19	0.03
0	900	0.92	0.87	-0.05	600	300	1.57	1.57	0.00	800	800	1.16	1.19	0.03
200	800	1.61	1.57	-0.04	1000	800	1.34	1.33	0.00	100	200	1.55	1.58	0.03
100	0	1.82	1.78	-0.04	900	500	1.55	1.55	0.00	0	600	1.22	1.26	0.03
300	0	1.82	1.79	-0.03	100	700	1.50	1.50	0.00	700	700	1.13	1.18	0.04
900	0	1.97	1.94	-0.03	900	300	1.72	1.72	0.00	300	500	1.18	1.23	0.05
600	1000	1.25	1.22	-0.03	200	200	1.57	1.57	0.00	200	600	1.39	1.44	0.05
0	300	1.45	1.42	-0.03	500	600	1.22	1.22	0.00	400	500	1.19	1.25	0.05
400	0	1.84	1.81	-0.03	600	100	1.78	1.78	0.00	800	400	1.53	1.58	0.05
900	700	1.39	1.36	-0.03	900	200	1.79	1.79	0.00	600	900	1.34	1.40	0.06
900	1000	1.16	1.14	-0.03	1000	300	1.76	1.76	0.00	100	900	1.01	1.09	0.08
500	0	1.86	1.83	-0.02	1000	100	1.92	1.92	0.00	1000	1000	1.10	1.19	0.09
900	400	1.65	1.63	-0.02	700	300	1.61	1.61	0.00	300	600	1.34	1.43	0.10
1000	900	1.22	1.20	-0.02	300	200	1.59	1.59	0.00	600	600	1.13	1.23	0.10
900	600	1.48	1.46	-0.02	500	300	1.53	1.53	0.00	600	700	1.39	1.49	0.10
100	1000	0.88	0.86	-0.02	100	600	1.39	1.39	0.00	200	500	1.13	1.24	0.10
200	1000	1.03	1.01	-0.02	100	100	1.69	1.69	0.00	700	900	1.25	1.37	0.11
600	0	1.87	1.85	-0.02	0	800	1.10	1.10	0.00	700	800	1.25	1.39	0.13
300	300	1.48	1.47	-0.02	800	1000	1.19	1.20	0.00	300	700	1.61	1.74	0.13
700	600	1.25	1.24	-0.01	400	100	1.72	1.73	0.00	0	500	1.10	1.24	0.14
600	500	1.31	1.30	-0.01	500	500	1.25	1.26	0.01	100	500	1.10	1.25	0.15
500	200	1.65	1.64	-0.01	700	200	1.70	1.71	0.01	400	600	1.25	1.49	0.24
1000	500	1.61	1.60	-0.01	1000	400	1.69	1.69	0.01	200	900	1.25	1.49	0.24
800	300	1.67	1.66	-0.01	900	900	1.16	1.17	0.01	0	1000	0.69	0.95	0.26
700	0	1.89	1.88	-0.01	400	200	1.61	1.62	0.01	400	900	1.61	1.96	0.35
800	600	1.39	1.38	-0.01	600	200	1.67	1.68	0.01	500	1000	1.28	1.79	0.51
800	500	1.48	1.47	-0.01	900	100	1.87	1.88	0.01	300	1000	1.25	1.81	0.55
400	300	1.50	1.50	-0.01	200	400	1.28	1.29	0.01	300	800	1.10	1.89	0.79
										500	800	0.69	1.89	1.20

Rezultatele numerice sintetizate în **tabelul 1.28** pot fi interpretate prin intermediul *parametrilor statistici descriptivi* și ai *reprezentărilor grafice*.

Parametrii statistici descriptivi semnificativi pentru corectitudinea modelului de variogramă se referă la diferența dintre grosimile măsurate (G_i) și cele estimate prin *kriging ordinar punctual* pe baza modelului de variogramă care se validează (G_i^*) și sunt:

- media erorilor de estimare = 0,001;
- abaterea standard a erorilor de estimare = 0,056;
- coeficientul de asimetrie al erorilor de estimare = 0,12.

Din valorile acestor parametri rezultă că valoarea erorilor de estimare reprezintă sub 0,1% din media valorilor grosimilor iar estimare este nedeviată (coeficientul de asimetrie indicând o distribuție normală a erorilor de estimare).

Reprezentările grafice utilizate pentru ilustrarea gradului de precizie al modelului de variogramă sunt:

- *Harta erorilor de estimare* care ilustrează distribuția spațială a erorilor de estimare (**Fig.1.75**).

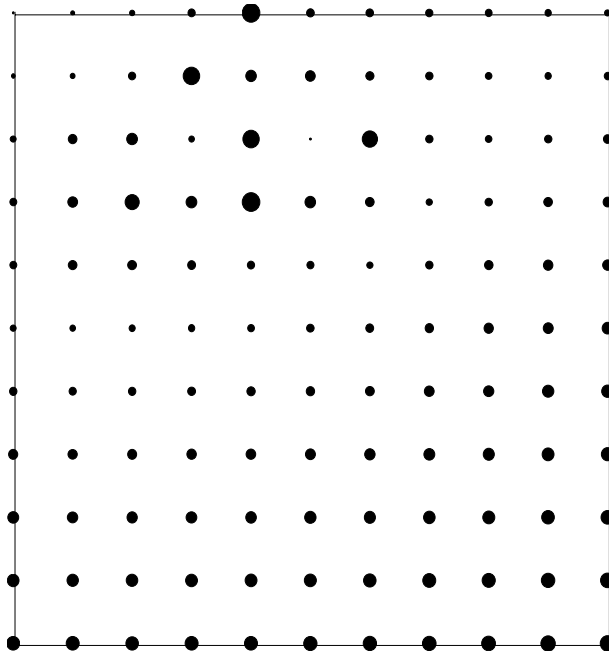


Fig.1.75. Harta erorilor de estimare

Pe această hartă erorile sunt reprezentate prin discuri negre al căror diametru este proporțional cu valoarea erorii de estimare. Se remarcă o grupare a erorilor mici în partea de NV a perimetrului explorat prin cele 121 de foraje și o creștere a erorilor de estimare spre SSE. De la această regulă se abat 6 locații în partea de N unde erorile sunt mai mari. Din această analiză sumară a hărții erorilor de estimare rezultă că modelul este “mai bun” pentru partea nordică decât pentru cea sudică a perimetrului.

De regulă în astfel de situații, pentru zona în care erorile sunt mai mari, se elaborează un alt model de variogramă pe baza valorilor măsurate în acea zonă dacă acestea sunt suficient de numeroase pentru a permite elaborarea unui model. Luarea unei astfel de decizii se face și în funcție de mărimea erorilor de estimare.

- *Diagrama grosime măsurată (G_i) - grosime calculată (G_i^*)* (**Fig.1.76**) indică o bună corelație a celor două categorii de grosimi. Dacă modelul de variogramă ar fi fost perfect, toate cele 121 de puncte din diagramă ar fi fost colineare. *Coeficientul de corelație lineară* dintre valorile măsurate (G_i) și cele calculate (G_i^*) cu ajutorul

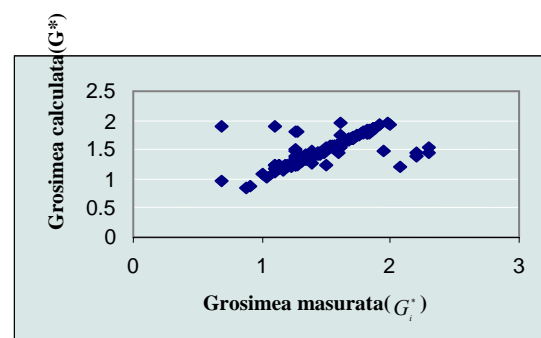


Fig.1.76. Diagrama $G_i - G_i^*$

modelului de variogramă testat este $r = 0,67$.

- *Diagrama grosime măsurată(G)- eroare de estimare (G-G*) (Fig.1.77) indică o estimare nedeviată a valorilor măsurate, prin dispunerea simetrică a erorilor pozitive și negative față de valoarea zero. Acest lucru este confirmat și de valoarea mediei erorilor de estimare (0,001) și a coeficientului de asimetrie a erorilor de estimare (0,12).*

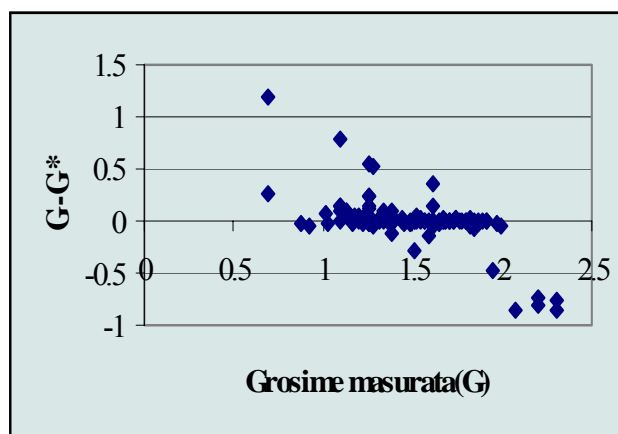


Fig.1.77. Diagrama grosime măsurată(G)- eroare de estimare (G-G*).

Reprezentările grafice descriu în mod diferit calitatea modelului de variogramă: *harta erorilor de estimare* indică distribuția în spațiul cercetat a valorii erorilor de estimare, iar *diagramele*, ignorând distribuția spațială a erorilor de estimare, evidențiază corelația între valoarea caracteristicii și mărimea erorii de estimare introdusă de modelul de variogramă.

COMENTARIU

Opțiunea pentru modelul testat se face în funcție de mărimea și distribuția erorilor de estimare raportate la un standard impus de criteriile unei estimări "eficiente". Eficiența estimării poate fi evaluată în raport cu gradul de cunoaștere al unei structuri spațiale sau cu riscul economic al unei investiții care se bazează pe estimările spațiale realizate cu ajutorul modelului.

Alegerea modelului optim impune validarea a cel puțin două modele !

Proportional cu creșterea suprafeței explorate scade probabilitatea ca un singur model să fie adecvat estimării spațiale pentru întreaga suprafață. În astfel de situații este necesar calculul variogramelor relative și utilizarea a două sau mai multe modele de variogramă pentru estimarea distribuției spațiale a caracteristicii pe întreaga suprafață cercetată.

Pentru această situație se apelează în prima fază la variogramele relative care au ca obiectiv ameliorarea morfologiei legii de variație spațială prin reducerea efectului distribuției neuniforme a punctelor de probare și a variațiilor locale determinate de valori extreme care nu au fost eliminate în faza analizei variabilității globale.

Utilizarea variogramelor relative poate conduce la două rezultate distincte, în funcție de caracteristicile structurii analizate:

- se identifică un singur model de variogramă valabil pentru întreaga suprafață studiată în etapa estimării distribuției spațiale;
- se identifică mai multe modele de variogramă valabile pentru diferite sectoare ale suprafeței studiate care vor fi utilizate în etapa de estimare spațială.

1.3.5. ANALIZA FRACTALĂ

Analiza structurilor spațiale de o mare complexitate, generate de procesele naturale, constituie obiectul unei metodologii relativ recente, al cărei act de naștere l-a semnat lucrarea revoluționară a matematicianului Benoit Mandelbrot, „*O teorie a seriilor fractale*”, apărută în anul 1975. Mandelbrot a introdus termenul de *fractal* pentru fenomene temporale sau spațiale care sunt *continui* dar *nediferențiabile*. *Fractalii* sunt primele forme grafice care nu sunt bazate pe *linii drepte* sau *liniaritate* și pot fi descrise prin dimensiuni fracționare.

Ideea de bază a calculului care a dominat gândirea științifică până la apariția fractalilor a fost aceea a aproximării unei *curbe* cu *linii drepte* și a unei *suprafețe* cu *plane*. În această idee, *curbele* și *suprafețele* pot fi reprezentate prin *funcții matematice continue* și *diferențiabile* pentru că pot fi împărțite într-un număr infinit de *linii drepte* sau *plane*. Leibnitz a presupus că toate curbele sunt alcătuite din *segmente de dreaptă infinitezimale* denumite *linii tangente* sau *diferențiale*. Orice concepție asupra Universului care utilizează teoriile de *diferențiere* și *integrare* se bazează pe această abordare *liniară*. Orice curs introductiv de calcul diferențial adoptă această idee prin utilizarea notației dh/dx care exprimă panta segmentului de dreaptă infinitezimal.

Conceptul de *fractal* se referă expres la structuri spațiale cu o dezordine accentuată, dezordine tratată ca o caracteristică intrinsecă a acestora și nu ca o perturbație. Caracteristica unei *structuri fractale* este conservarea sau repetarea dezordinii la toate scările, repetarea structurilor în structuri (autosimilaritatea).

În funcție de tipul și gradul de dezordine se pot separa:

- structuri spațiale cu o *dezordine slabă*, pentru care dezordinea dispare sau se reduce pe măsură ce crește sau se reduce scara de investigare;
- structuri spațiale cu o *dezordine accentuată*, pentru care dezordinea persistă pe măsură ce scara de investigare crește sau se reduce și nu există stări de referință cu care să poată fi comparat sistemul.

Dezordinea structurii spațiale este cuantificată de *dimensiunea fractală* (D); cu cât complexitatea structurii spațiale este mai mare cu atât D este mai mare. *Analiza fractală* abordează studiul structurilor spațiale renunțând la „*primordialitatea liniarității*” utilizând *dimensiunea fractală* (D) și *principiul autosimilarității*.

Dimensiunea fractală este definită pe baza a două modele diferite de abordare a structurilor spațiale. Unele dintre acestea reprezintă structura ca un *model continuu* ce poate fi investigat la o scară infinit mică, iar altele presupun reprezentarea acesteia printr-un *model discontinuu*, *tip rețea*.

Determinarea *dimensiunii fractale* pentru un *model continuu*, de tipul unei *curbe* sau *suprafețe* neregulate, presupune împărțirea spațiului în care se dezvoltă structura studiată printr-o *rețea pătratică* de parametru ε și numărarea celulelor pătrate care intersectează contururile acesteia.

Numărul celulelor care intersectează structura studiată ($N_{box}(\varepsilon)$) poate fi utilizat pentru realizarea estimărilor de tipul:

- *lungimea conturului curbei* (l):

$$l = N_{box}(\varepsilon) \cdot \varepsilon \quad (1.73)$$

- aria delimitată de conturul neregulat (A):

$$A = N_{box}(\varepsilon) \cdot \varepsilon^2 \quad (1.74)$$

Numărul celulelor care intersectează structura studiată este:

- pentru o *linie dreaptă*, proporțional cu parametrul rețelei pătratice (ε):

$$N_{box}(\varepsilon) \propto \varepsilon^{-1} \quad (1.75)$$

- pentru un *pătrat*, proporțional cu pătratul parametrului rețelei pătratice (ε):

$$N_{box}(\varepsilon) \propto \varepsilon^{-2} \quad (1.76)$$

- pentru o *linie neregulată*, proporționalitatea este dată de un coeficient D , care este *dimensiunea fractală* a structurii:

$$N_{box}(\varepsilon) \propto \varepsilon^{-D} \quad (1.77)$$

D coincide cu *dimensiunea topologică* (D_{top}) a structurii dacă sistemul este *euclidian*:

$D_{top} = 0$ pentru un set de puncte discrete;

$D_{top} = 1$ pentru o curbă;

$D_{top} = 2$ pentru o suprafață;

$D_{top} = 3$ pentru un solid.

Dacă D diferă de D_{top} , sistemul este numit *fractal*, D este fracționar și este *dimensiunea fractală*. Diferența $D_{top} - D$ este o măsură a dezordinii sistemului.

În sisteme reale dimensiunea celulelor pătratice, ε , este finită: ε_{min} se numește *limită internă* și teoretic merge până la dimensiunile atomice, iar ε_{max} se numește *limită externă* și poate ajunge până la distanța dintre cele mai îndepărtate puncte ale structurii spațiale. În general ε_{min} și ε_{max} nu coincid cu limitele teoretice pentru că, de exemplu, peste sau sub o anumită scară de cercetare structurile spațiale pot fi euclidiene și deci nu mai au dimensiune fractală.

Faptul că lungimea unei drepte care face cu rețeaua de discretizare un anumit unghi este aproximată cu lungimea unei linii “în trepte” (**Fig.1.78a**), rezultând o supraestimare, a condus la înlocuirea *pătratelor* cu latura ε cu *cercuri* de rază r (**Fig.1.78b**). În aceste condiții lungimea dreptei este corect estimată prin relația:

$$l = N_{cerc}(r) \cdot 2 \cdot r \quad (1.78)$$

Indiferent că “acoperim” structura spațială cu pătrate sau cu cercuri, *dimensiunea fractală* D nu se modifică. Forma elementului care este utilizat pentru “acoperire” afectează numai factorul de proporționalitate (\propto) dintre numărul de elemente (N_{box}/N_{cerc}) și dimensiunea acestora (ε/r).

Calculul dimensiunii D a unei structuri spațiale se face cu relația:

$$D = - \frac{\ln(N(\varepsilon))}{\ln(\lambda)} \quad (1.79)$$

în care

$N(\varepsilon)$ este numărul de elemente de dimensiune ε care acoperă sistemul;

λ - raportul dintre ε și dimensiunea maximă a sistemului.

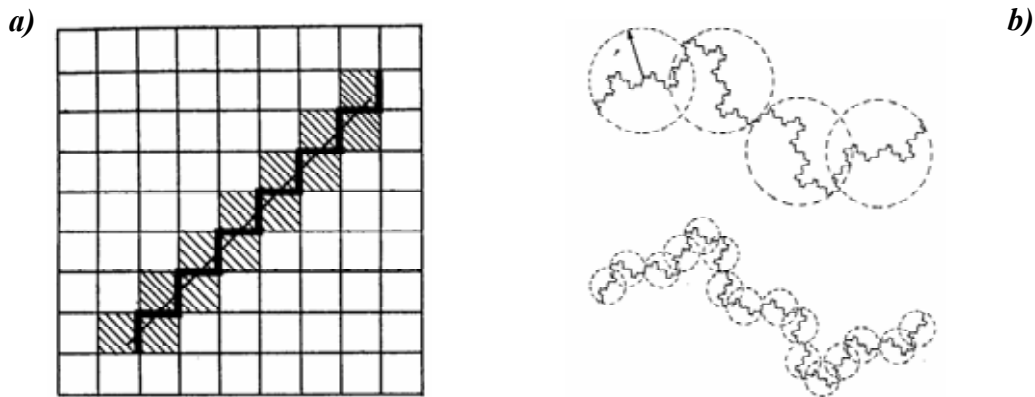


Fig.1.78. Aproximarea lungimii unei drepte care face un unghi cu rețeaua de discretizare pătratică (a) și a unei curbe prin acoperire cu cercuri (b).

Relația (1.79) este valabilă și pentru *structurile euclidiene*, rezultatul fiind *dimensiunile topologice*. Astfel:

- un segment rectiliniu de lungime 1 m este “acoperit” cu $N(0,25) = 4$ pătrate cu latura 0,25 m și are dimensiunea topologică:

$$D = -\frac{\ln(4)}{\ln(0,25/1)} = \frac{\ln(4)}{\ln(1/4)} = 1$$

- un pătrat cu latura de 10 m este “acoperit” cu $N(2) = 25$ de pătrate cu latura egală cu 2 m și are dimensiunea topologică:

$$D = -\frac{\ln(25)}{\ln(2/10)} = -\frac{\ln(25)}{\ln(1/5)} = 2$$

- un cub cu latura 10 m este divizat în $N(5) = 8$ cuburi cu latura de 5 m și are dimensiunea topologică:

$$D = -\frac{\ln(8)}{\ln(5/10)} = -\frac{\ln(8)}{\ln(1/2)} = 3$$

Reluând într-o exprimare generalizantă, putem spune că dacă un obiect cu o formă dată este compus din ε^D obiecte similare de mărimea $1/\varepsilon$ atunci exponentul D este *dimensiunea de similaritate* și pentru că, spre deosebire de cea topologică, această dimensiune *nu este în mod obligatoriu întregă*, ea este numită și *dimensiune fractală*.

Pentru structurile spațiale fractale, deoarece acoperirea nu este completă pentru orice mărime a elementelor utilizate, se calculează *dimensiunea fractală medie* pentru mai multe mărimi ale elementelor de acoperire.

Modelul discontinuu este utilizat pentru sisteme cu evidente proprietăți de *autosimilaritate*. Principiul de bază al modelului discontinuu tip rețea, pentru evaluarea dimensiunii fractale D , constă în numărarea locațiilor din structura spațială pentru o anumită vecinătate (R) (Fig.1.79). Locațiile din structura spațială sunt puncte plasate la o anumită distanță între ele (d). Numărul locațiilor se numără într-o sferă de rază $R \gg d$ și este proporțional cu dimensiunea D a structurii spațiale:

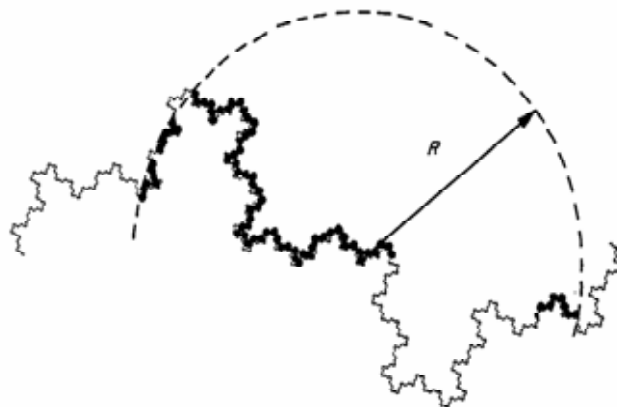


Fig.1.79. Modelul discontinuu tip rețea

$$N_{loc}(R) \propto R^D \quad (1.80)$$

Dacă structura spațială rezultă din combinarea a m copii identice la scară mai mică, de factor λ , atunci:

$$N_{loc}(R/\lambda) = m \cdot N_{loc}(R) \quad (1.81)$$

Exprimând R în multipli de d se poate scrie relația pe baza căreia se calculează dimensiunea fractală pentru modelul discontinuu tip rețea:

$$N_{loc}(R) = B \cdot \left(\frac{R}{d}\right)^D \quad (1.82)$$

Factorul B este mic dacă centrul sferei de rază R este într-o zonă fără locații. Dacă fluctuațiile lui B sunt mici sistemul este de *lacunaritate* scăzută, caracteristică sistemelor euclidiene. Pentru o dreaptă și un plan fluctuațiile lui B sunt nule și *lacunaritatea* este zero.

Variograma, instrumentul de bază al analizei variografice permite și ea determinarea dimensiunii fractale a structurilor spațiale în condițiile bunei cunoașteri a acestora prin rețeaua de explorare.

Există o fundamentare logică a acestei posibilități pentru că dimensiunea fractală este o cuantificare a complexității structurii spațiale, iar variograma cuantifică variabilitatea spațială a structurilor în funcție de distanță și direcție. La o analiză elementară se poate observa dependența dintre complexitatea structurii spațiale, forma variogramei și valoarea dimensiunii fractale (Fig.1.80). Sunt remarcabile:

- proporționalitatea directă dintre complexitatea structurii și dimensiunea fractală;
- invers proporționalitatea dintre complexitatea structurii și panta variogramei.

Adresându-se structurilor spațiale complexe, a căror dezordine o consideră caracteristica intrinsecă, evaluarea dimensiunii fractale trebuie să se bazeze pe o "informație care să conțină această dezordine". De aici rezultă necesitatea unui număr

mare de puncte de observație pentru calculul unei variograme care să permită evaluarea corectă a dimensiunii fractale.

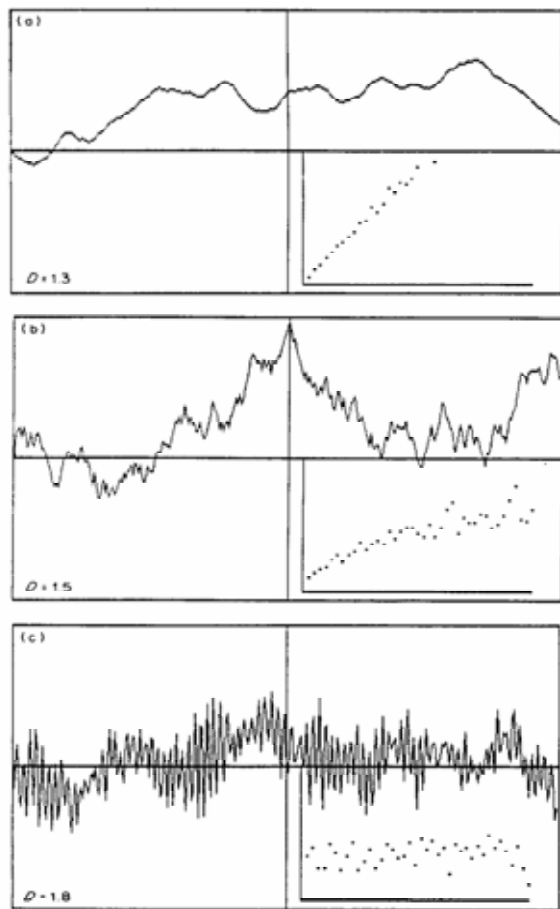


Fig.1.80. Corelația între complexitatea structurii, dimensiunea fractală și variogramă.

a avea aceleași caracteristici morfologice indiferent de scara la care este examinată sau, altfel spus, se poate descompune în componente geometrice similare cu întreaga structură.

Multe fenomene naturale par să aibă această caracteristică. Dacă distribuția litotipurilor într-o zonă este cartografiată la scara 1:1.000.000 și ulterior se avansează cu detalierea la scări mai mari, de exemplu 1:100.000, distribuția relativă a limitelor, deseori, pare a nu fi influențată de scara la care este construită harta.

Acest fenomen de *autosimilaritate* poate fi cuantificat prin *probabilitatea de apariție a limitelor* dintre litotipi (p) la o anumită *distanță* (d) probabilitate care este exprimată satisfăcător de *funcția de distribuție exponențială*:

$$p(d) = 1 - e^{-\alpha d} \quad (1.83)$$

Valoarea probabilității, $p(d)$, este cuprinsă între 0 și 1, d indică distanța dintre limite iar α este un parametru de scară. În multe situații distribuția cumulativă a limitelor dintre litotipi este mai bine descrisă de o *distribuție gamma sau Weibull* dar, în limitele unei precizii satisfăcătoare, ele pot fi reduse la cea *exponențială*.

În toate situațiile, dimensiunea fractală se calculează din panta variogramei într-un sistem de referință dublu logaritm ($\log(\gamma(h)) - \log(h)$); $\gamma(h)$ - valoarea semivariogramei; h - distanța dintre punctele de observație).

Foarte frecvent, într-un sistem dublu logaritm, variogramele sunt lineare pe domeniul distanțelor dintre punctele de probare, acest lucru exprimând menținerea caracteristicilor de variabilitate spațială în arealul cercetat și demonstrând proprietatea de *autosimilaritate* a structurii spațiale. Modificări ale pantei variogramelor indică schimbarea legăturilor de variație spațială, corespunzător scării de investigare, adică a distanțelor dintre punctele de probare.

Principiul autosimilarității este a doua idee fundamentală care stă la baza analizei fractale. *Autosimilaritatea* este proprietatea unei *structuri spațiale fractale* de

A18. Calculul dimensiunii fractale a unei structuri de creștere

Să se determine dimensiunea fractală a structurii de blendă din Fig.1.81.

Rezolvare

Structura analizată are patru brațe care nu se intersectează într-un punct comun ci sunt legate printr-o "punte" care se constituie într-o a treia direcție de creștere. Ramura din dreapta jos se bifurcă, iar unul din brațele sale se întoarce la 180 grade (Fig.1.81).



Fig.1.81. Structură de blendă

Structura s-a produs prin difuzie în mediu solid și agregare. Ea este rezultatul dezamestecului sulfurii de zinc (blenda) de sulfura de cupru (calcopirita) dintr-o soluție solidă, instabilă la temperaturi scăzute. Proba studiată (E.Constantinescu, S.Milutinovici, 1994) a fost recoltată din zăcamântul Toroiaga din zona vulcanitelor neogene.

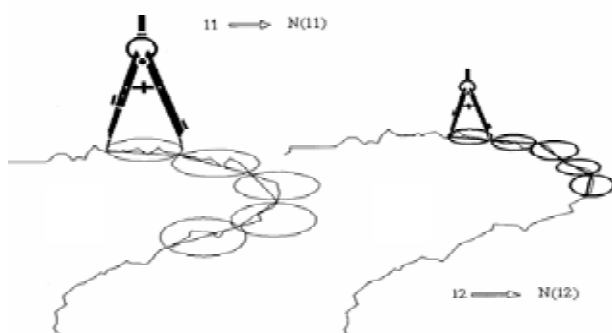


Fig.1.82 Calculul dimensiunii fractale prin metoda perimetrului.

lungime l_i se obține un număr N_i de segmente cu care se acoperă conturul structurii. Cu cât l_i este mai mic cu atât N_i este mai mare și dacă structura este un obiect fractal, relația dintre N_i și l_i va fi de forma:

$$N_i = l_i^{-D} \quad (1.84)$$

în care D este dimensiunea fractală.

În general se utilizează o metodă grafică în estimarea dimensiunii fractale, metodă care presupune linearizarea corelației dintre N_i și l_i într-o diagramă rectangulară dublu logaritmică.

Metoda utilizată pentru determinarea dimensiunii fractale a structurii este cea a măsurării perimetrului. Ea constă în aproximarea prin segmente de dreaptă a lungimii structurii studiate. Se alege unul din capetele conturului (oricare) ca punct de start și folosind un compas se marchează intersecțiile succesive ale arcelor de cerc cu conturul, utilizând fiecare punct astfel obținut ca punct de start (Fig.1.82). Pentru fiecare

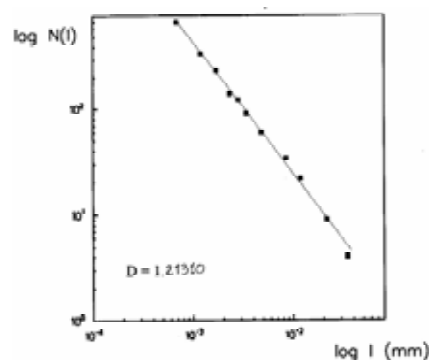


Fig.1.83. Calculul dimensiunii fractale prin metoda grafică

Dimensiunea fractală (D) reprezintă panta drepte de corelație. Pentru structura de creștere a sulfurii de zinc, dimensiunea obținută prin această metodă grafică (**Fig.1.83**) este $D = 1,2131$.

COMENTARIU

Metoda măsurării perimetrului este o metodă care nu poate fi automatizată, motiv pentru care ea este utilizată foarte rar. Parcurgerea acestei metode este însă una din căile de acomodare cu noțiunea de dimensiune fractală.

În general, dacă nu cunoaștem și nu înțelegem în detaliu toate caracteristicile unei noțiuni, nu vom putea opera corect cu ea.

Dimensiunea fractală este o noțiune fundamentală a analizei fractale prin care se abordează studiul formelor din spațiul tridimensional în care trăim de pe alte fundamente decât cele cu care s-a operat până acum 25 de ani.

Rezistăm la această provocare?

După surpriza inițială a noii abordări a urmat un impresionant asalt al aplicării acesteia în toate domeniile științei. Entuziasmul s-a redus însă exponențial, atunci când câștigul s-a dovedit a fi nesemnificativ în studiul formelor reale.

A19. Analiza distribuției fractale

Să se compare condițiile de formare ale zăcămintelor Masvingo-Mashava și Shamva din cratonul arhaean (din Zimbabwe, **Fig.1.84** și **Fig.1.85**).

Rezolvare:

Analiza distribuției spațiale a zăcămintelor poate evidenția caracteristici mascate de complexitatea proceselor de formare a acestora, caracteristici care pot fi utilizate pentru explorarea și evaluarea rezervelor.

Dacă zăcămintele au o distribuție fractală, ele sunt grupate spațial într-un anumit fel și pot fi tratate cu instrumentele specifice geometriei fractale. Analiza dimensionalității structurilor poate discrimina cu precizie distribuțiile aleatoare de cele fractale.

Pentru exemplificare sunt prezentate rezultatele obținute printr-o analiză fractală completă asupra celor două zăcămintele de Carlson și T.Blenkinshop (1991; 1994). Localizarea zăcămintelor din cele două arii și statistica producției acestora sunt obținute din buletinele și hărțile la scara 1:100.000 elaborate de Geological Survey of Zimbabwe (Wilson 1964, 1968; Stidloph, 1977).

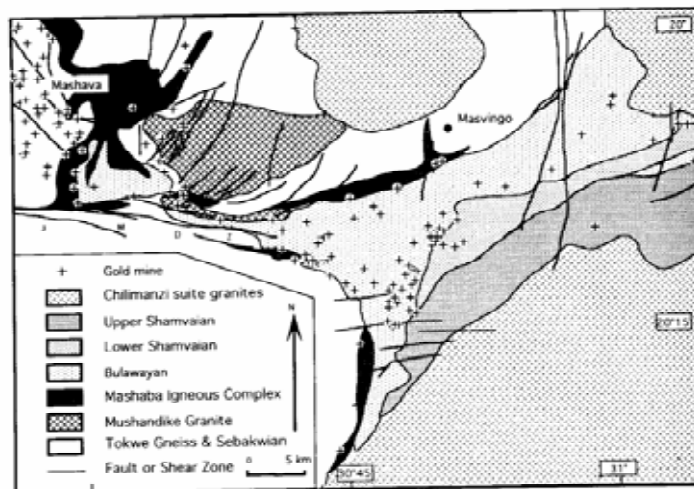


Fig.1.84. Zona Masvingo-Mashava
(după T.Blenkinshop, 1994)

Zona Masvingo-Mashava a produs aproximativ 7000 kg aur în perioada

studiată, din 147 de mine care pot fi bine localizate (Fig.1.84) și care au fost utilizate pentru realizarea analizei fractale. Zona Shamva a produs 56874 kg de aur în aceeași perioadă, din 122 de mine (Fig.1.85).

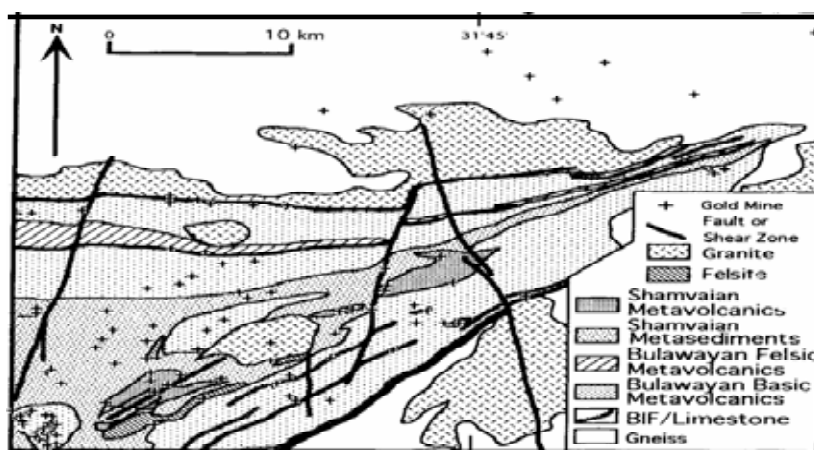


Fig.1.85. Zona Shamva (după T.Blenkinshop, 1994)

Au fost utilizate două metode pentru calculul dimensiunii fractale:

- metoda numărării celulelor, care constă în numărarea celulelor pătrate ($n(d)$) de o anumită dimensiune (d) necesare acoperirii tuturor minelor din zona studiată:

$$n(d) \propto d^{-D} \quad (1.85)$$

- metoda numărului din cerc, care constă în numărarea zăcămintelor ($N(r)$) dintr-un cerc de rază r :

$$N(r) = C \cdot r^D \quad (1.86)$$

în care

D - dimensiunea fractală a structurii;

C - o constantă numerică adimensională.

S-a demonstrat (Carlson, 1991) că densitatea zăcămintelor într-un cerc de rază r ($\rho(r)$) și densitatea depozitelor $d(r)$ la distanța r pot fi deduse sub forma:

$$\rho(r) = \frac{C}{\pi \cdot r^{D-2}} \text{ și } d(r) = \frac{D \cdot C}{2 \cdot \pi \cdot r^{D-2}} = \frac{D}{2 \cdot \rho(r)} \quad (1.87)$$

$N(r)$ este evaluat pentru fiecare zăcământ în parte iar rezultatele pentru întreaga zonă sunt mediate pentru a obține relația finală.

D este evaluat prin regresie lineară iar erorile sunt exprimate prin abaterea standard a acesteia.

Pentru ambele metode, dimensiunile lineare d și r variază de la o valoare minimă, mai mare decât eroarea de măsurare a distanțelor ($d = 300 \text{ m}$; $r = 150 \text{ m}$), la una maximă, de ordinul de mărime al zonei studiate ($d = 29 \text{ km}$; $r = 23 \text{ km}$).

Graficele $n(d) - d$ (Fig.1.86a) și $N(r) - r$ (Fig.1.86b) indică pentru ambele zone studiate forme similare.

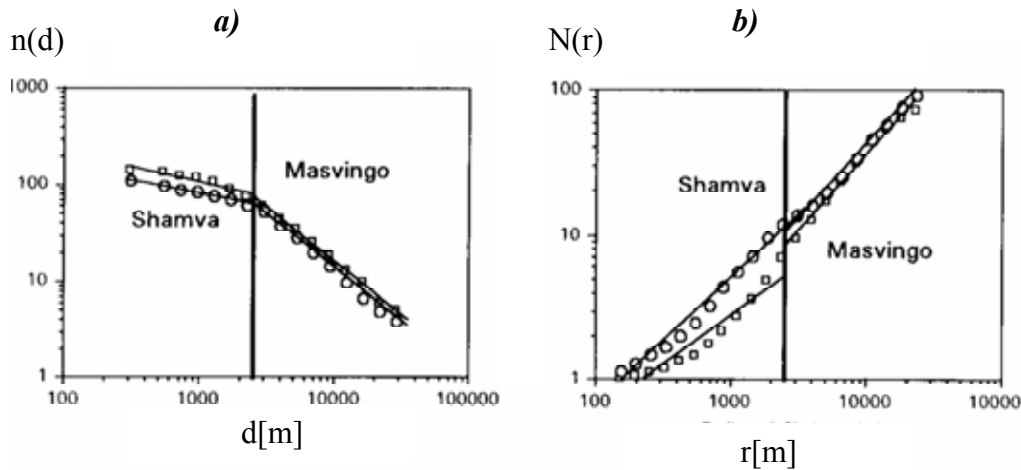


Fig.1.86. Estimarea dimensiunii fractale prin metoda numărării celulelor (a) și metoda numărului din cerc (b)

Metoda numărării celulelor indică o modificare semnificativă a dimensionalității fractale la o valoare $d = 2,5$ km. La distanțe mai mari de 2,5 km manifestarea fractalității structurii este mai evidentă și o detaliere a diagramelor pentru acest domeniu este ilustrativă (Fig.1.87a și b). Parametrii statistici asociați

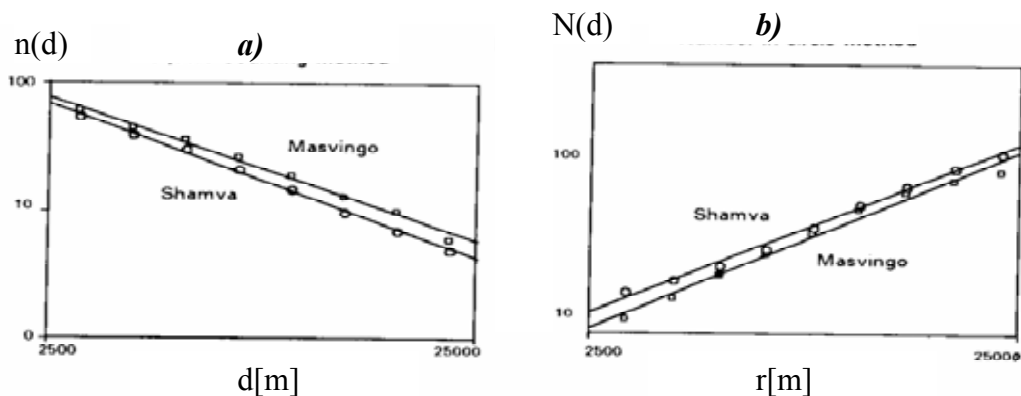


Fig.1.87. Estimarea dimensiunii fractale prin metoda numărării celulelor (a) și metoda numărului din cerc (b) pentru distanțe mai mari de 2500 m.

determinării dimensiunii fractale prin metoda numărării celulelor sunt sintetizați în tabelul 1.29 iar pentru metoda numărului din cerc în tabelul 1.30.

Rezultatele arată o serie de diferențe între cele două metode utilizate.

Metoda numărării celulelor conduce la valori mai mici pentru D în domeniul valorilor mici ale lui d sau r și valori mai mari pentru D în domeniul valorilor mari pentru d sau r , decât metoda numărului din cerc. Acest lucru se întâmplă deoarece metoda numărării celulelor este mai sensibilă la corelația între dimensiunea fractală și scara de investigare, indicând o schimbare semnificativă a relației fractale la $d = 2500$ m. Dimensiunea fractală pentru d sau r mai mari de 2500 m, obținută prin metoda numărării celulelor, este mai mare decât cea determinată prin metoda numărului din cerc. Erorile de determinare ale dimensiunii fractale coincid pentru cele două metode în zona Masvingo-Mashava dar sunt diferite pentru zona Shamva. Metoda numărării celulelor conduce la un coeficient de corelație mai mare și la erori ale regresiei mai mici decât metoda numărului din cerc.

Tabelul 1.29. *Determinarea dimensiunii fractale prin metoda numărării celulelor.*

Zona	Distanța (d) [km]	Dimensiunea fractală (D)	Coefficientul regresiei lineare (R)	Abaterea standard a regresiei (s)	Nr. De zăc. (N)
Masvingo	0,3-2,5	0,32	-0,947	0,0415	147
Shamva	0,3-2,5	0,28	-0,988	0,0169	122
Masvingo	2,5-29	1,12	-0,995	0,0379	147
Shamva	2,5-2,9	1,18	-0,999	0,0192	122

Tabelul 1.30. *Determinarea dimensiunii fractale prin metoda numărului din cerc*

Zona	Distanța (r) [km]	Dimensiunea fractală (D)	Coefficientul regresiei lineare (R)	Abaterea standard a regresiei (s)	Nr. de zăc. (N)
Masvingo	0,15-2,5	0,69	-0,958	0,0595	147
Shamva	0,15-2,5	0,87	-0,991	0,0347	122
Masvingo	2,5-22,5	1,06	-0,994	0,0400	147
Shamva	2,5-22,5	1,01	-0,997	0,0246	122

Toate aceste observații sugerează că *metoda numărării celulelor* este mai adecvată pentru descrierea detaliată a distribuției fractale pentru un domeniu mai extins al scării de investigare.

Interpretarea acestor rezultate servește *comparării zăcămintelor* din cele două zone. Se constată că pentru ambele zone se produce o modificare a relațiilor fractale la distanțe de investigare de 2500 m. Acest lucru sugerează că este o mai mare densitate a zăcămintelor la distanțe mai mici decât această distanță ceea ce este în concordanță cu concentrarea efortului de explorare în vecinătatea zonelor mineralizate.

Adevărata dimensiune fractală poate fi stabilită numai la dimensiuni mai mari ale scării de investigare și obținerea în aceste condiții a aceleiași valori pentru ambele zone sugerează că aceleași procese geologice au generat zăcămintele din ambele zone cercetate.

Această abordare se bazează numai pe distribuția spațială a zăcămintelor. Luarea în considerare și a distribuției spațiale a conținuturilor, într-o *analiză multifractală*, oferă date suplimentare pentru analiza comparativă a zăcămintelor din cele două zone.

COMENTARIU

Analiza fractală ca și cea variografică are ca obiectiv cuantificarea caracteristicilor esențiale ale structurii spațiale cercetate.

Analiza variografică are ca produs final modelul de variogramă cu parametrii caracteristici (efect de pepită, palier, rază de influență, forma analitică), iar analiza fractală are ca produs dimensiunea fractală (echivalentă într-o anumită măsură cu raza de influență a modelului de variogramă).

Corectitudinea modelului de variogramă și a dimensiunii fractale este decisivă pentru precizia estimării distribuției spațiale.

2. ESTIMAREA DISTRIBUȚIEI SPAȚIALE

Estimarea distribuției spațiale fructifică rezultatele analizei variabilității caracteristicilor studiate. Ea echivalează cu realizarea *desenului* după analiza obiectului supus studiului.

Calitatea *desenului* este determinată de calitatea stimulilor (numărul de puncte de observație și precizia determinărilor în cazul cercetării geologice), de profunzimea analizei, de sensibilitatea și experiența creatorului. Nu în ultimul rând, calitatea desenului depinde de materialul utilizat (creion, hârtie) și de acuratețea cu care mâna răspunde comenzii creatorului.

S-ar putea ca pe parcursul realizării *desenului* (conturilor hărții sau secțiunii geologice) să constăți că imaginea creată nu este ce ți-ai dorit sau nu-ți place. Câte tablouri zac acoperite de praful uitării sau nici nu au ieșit din atelierul creatorului?

Dacă nu ești mulțumit de rezultat arunci la coș tot ce-ai făcut și o iei de la capăt. Șansele ca *desenul* să fie mai bun a doua oară sunt mai mari. Nu este sigur că după a mia încercare rezultatul este de o mie de ori mai bun decât primul. *Nu toți reușesc în viață, dar merită să încerci!*

Estimarea distribuției spațiale este o operațiune laborioasă și rezultatul ei este totdeauna discutabil. Obiectivul ei este realizarea *imaginii* distribuției spațiale a unei variabile *pe baza valorilor acelei variabile determinate în diferite puncte din spațiu*. Pentru geolog aceste *imagini* realizate prin *metode geostatistice* sunt *hărțile* și *secțiunile geologice*.

Ambele categorii de reprezentări grafice se realizează atât pentru variabile *calitative* cât și pentru cele *cantitative*:

- hărțile și secțiunile geologice clasice reprezintă distribuția variabilelor de tip *calitativ* precum *vârsta* sau *litologia* formațiunilor;
- hărțile și secțiunile speciale ilustrează variația spațială a diferitelor caracteristici *cantitative*: adâncimea sau cota unui reper stratigrafic sau structural (*hărți cu izobate sau izohipse*), grosimea unor depozite cu o anumită litologie (*hărți cu izocore*), potențialul unor acvifere (*hărți cu izotransmisivități*), accelerația gravitațională într-o zonă prospectată gravimetric (*hărți cu izogale*), intensitatea activității seismice (*hărți cu izoseiste*) etc.

Hărțile și secțiunile geologice sunt în același timp suportul și rezultatul cercetărilor geologice. Ele sunt folosite pentru clarificarea structurilor geologice, identificarea zăcămintelor, calculul rezervelor de substanțe minerale utile, protecția mediului. Tehnica de realizare a hărții sau secțiunii se alege în funcție de obiectivul pentru care este utilizată. Într-o clasificare simplă se pot separa trei categorii de astfel de tehnici:

- *estimarea globală*, prin care se obține *valoarea medie reprezentativă* a unei variabile pentru *toată suprafața hărții*. Metoda este utilizată de obicei în faza de prospecțiune pentru evaluarea aproximativă a rezervelor și se aplică în cazul unui număr redus de puncte de observație;
- *estimarea punctuală* este cea mai flexibilă și performantă tehnică de realizare a tuturor categoriilor de hărți cu izolinii. Ea dispune de numeroase variante de prelucrare geostatistică (*kriging punctual ordinar, k.p.universal, k.p.indicator, k.p.disjunctiv* etc.);
- *estimarea zonală* se utilizează pentru *calculul rezervelor* în etapa de explorare detaliată și permite vizualizarea distribuției conținuturilor de substanțe minerale utile pe suprafețe/volume delimitate de contururi cu geometrie cunoscută (ex.: panouri de exploatare, orizonturi miniere etc.).

După ce am colectat *probe* din *n puncte de observație* (puncte marcate cu *x-uri* în **Fig.2.1**), am determinat prin diferite metode *n valori* ale unei *variabile* ($v_i; i=1 \dots n$) și le-am *analizat variabilitatea (globală și spațială)*, avem la dispoziție informațiile și instrumentele necesare pentru a putea răspunde la oricare din următoarele întrebări:

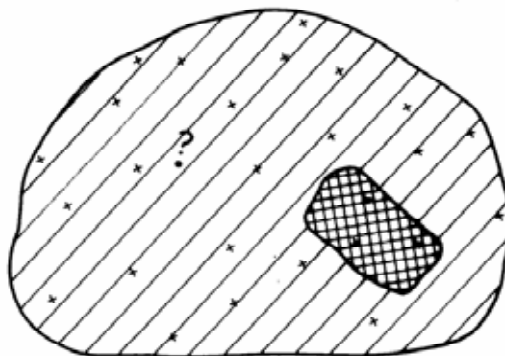


Fig. 2.1. Obiectivele estimării spațiale.

- Care este *valoarea medie a variabilei pe toată suprafața probată* (suprafața marcată cu hașură simplă din **Fig.2.1**)?
- Care este *valoarea medie a variabilei probate pe suprafețe delimitate de contururi închise, mai mici decât întreaga suprafață probată* (suprafața marcată prin hașură încrucișată în **Fig.2.1**)?
- Care este *valoarea variabilei într-un punct oarecare al suprafeței studiate* (de exemplu punctul marcat cu semnul întrebării în **Fig.2.1**, punct în care nu am avut acces pentru a măsura variabila).

Indiferent de întrebarea care se pune, răspunsul îl dau metodele *estimării distribuției spațiale univariate*, metode care se bazează de cele mai multe ori pe o combinație lineară de forma:

$$v_o^* = \sum_{i=1}^n w_i v_i \quad (2.1)$$

în care:

- v_o^* - valoarea estimată a variabilei într-un punct oarecare p_0 ;
- v_i - valorile cunoscute ale variabilei în punctele p_i ($i=1, 2, \dots, n$);
- w_i - ponderile acordate fiecărei valori măsurate.

Combi-națiile lineare de tipul relației (2.1) operează și asupra datelor transformate (*normalize*). Dacă se transformă datele originale și se operează asupra lor o combinație lineară ponderată, se obține un estimator al valorilor transformate:

$$t_o^* = \sum_{i=1}^n w_i T(v_i) \quad (2.2)$$

în care:

- t_o^* - estimatorul valorii transformate a variabilei într-un punct oarecare p_0 ;
- $T(v)$ - funcția de transformare a valorilor variabilei v ;
- w_i - ponderile acordate fiecărei valori transformate.

În astfel de situații, pentru obținerea valorii estimate a variabilei (v_o^*) este necesară transformarea inversă a estimatorului valorilor transformate (t_o^*):

$$v_o^* = T^{-1}(t_o^*) \quad (2.3)$$

Obiectivul operațional principal al estimării distribuției spațiale este calculul ponderilor acordate valorilor măsurate (w_i) în estimarea realizată!

2.1. ESTIMAREA GLOBALĂ

Evaluarea *valorii medii reprezentative* pentru o *variabilă regionalizată staționară* pe toată suprafața probată, *medie* utilizată în estimările geostatistice bazate pe funcția de covarianță, implică eliminarea efectului distribuției spațiale neuniforme a punctelor de observație. Gruparea neuniformă a punctelor de observație poate conduce la *subestimarea* valorii *mediei*, dacă majoritatea punctelor de probare sunt concentrate în zone cu *valori mici* ale variabilei, sau la *supraestimări* ale acesteia, în caz contrar.

Eliminarea efectului distribuției neuniforme a punctelor de observație asupra valorii mediei globale se poate realiza prin *declustering poligonal* sau *celular* și prin *kriging*.

Declusteringul poligonal, pentru eliminarea influenței grupării neuniforme a punctelor de observație atribuie fiecărei valori din setul de date o pondere proporțională cu suprafața *poligonului de influență*.

Poligonul de influență pentru un punct de observație rezultă din intersecția mediatoarelor segmentelor ce unesc centrul poligonului de influență (O) cu punctele vecine (1,2,3,4,5,6). Punctul 7 nefiind în vecinătatea imediată a punctului O nu modifică forma poligonului de influență al acestuia (Fig.2.2a,...,f).

Definitivarea poligoanelor de influență pentru punctele din *vecinătatea limitelor ariei explorate* implică introducerea unor reguli suplimentare. Există două astfel de reguli: închiderea poligonului pe o *limită fizică* (Fig.2.3a) sau *convențională* a ariei explorate (Fig.2.3.b) conturată printr-un arc de cerc a cărui rază este media aritmetică a apotemelor OA, OB și OC.

Separarea în poligoane de influență pe baza metodei prezentate este *unică*.

Fiecărui punct din rețeaua de observație îi este asociat un poligon cu suprafața mai mare în zonele cu densitate mică de puncte de observație și cu suprafață mai mică în zonele cu densitate mai mare de puncte (Fig.2.4).

Pentru calculul ponderilor standardizate ale fiecărei valori se utilizează relația:

$$w_i = \frac{a_i}{\sum_{i=1}^n a_i}, \quad (2.4)$$

în care a_i este aria fiecărui poligon de influență ($i = 1, 2, \dots, n$; n - numărul punctelor de observație).

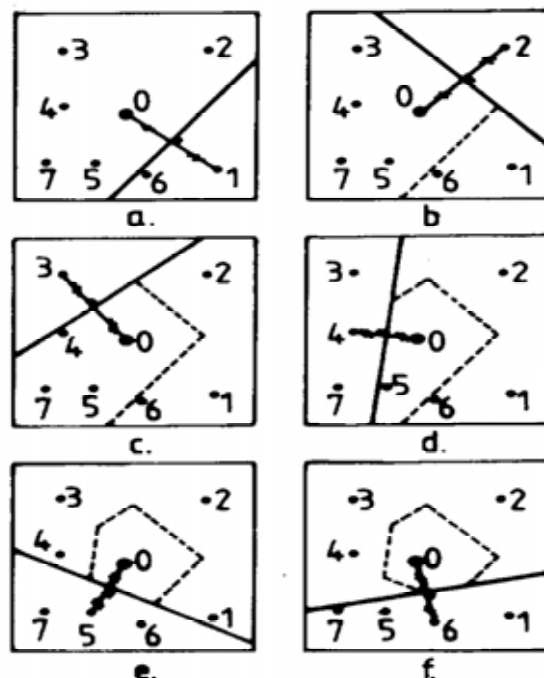


Fig. 2.2 Construcția poligonului de influență pentru un punct din interiorul suprafeței probate.

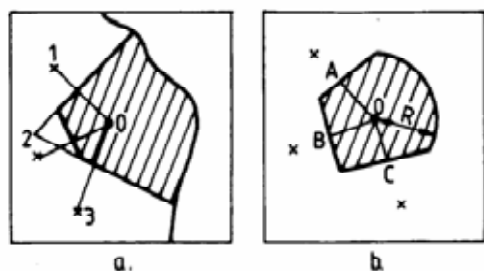


Fig. 2.3. Construirea poligonului de influență pentru un punct din vecinătatea limitei ariei explorate

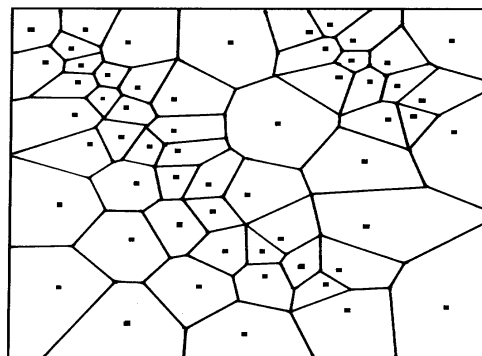


Fig. 2.4. Poligoanele de influență ale punctelor de observație dintr-o rețea neuniformă.

Declusteringul celular aplică pe întreaga suprafață a ariei explorate o rețea rectangulară. Fiecare valoare primește o pondere invers proporțională cu numărul de puncte din celula căreia îi aparține. În acest fel, valorile plasate în zone cu densitate mai mare a punctelor de observație primesc o pondere mai mică, iar valorile plasate în zonele cu densitate mai mică a punctelor de observație primesc o pondere mai mare.

Pentru dispoziția punctelor de probare din **Fig. 2.5**, plasate în celulele unei rețele rectangulare cu dimensiunea, pe direcția V-E, de 100 m și pe direcția N-S, de 50 m, ponderile acordate valorilor sunt:

- pentru celula A3 : $w = 1/4$;
- pentru celula A2 : $w = 1/6$;
- pentru celula A1 : $w = 1/2$;
- pentru celula B3 : $w = 1$;
- pentru celula B2 : $w = 1$;
- pentru celula B1 : $w = 1/5$.

Deoarece toate valorile din fiecare celulă au ponderi egale și ponderea valorilor din fiecare celulă este unitară, estimarea mediei globale prin metoda declusteringului celular se realizează în două etape:

- a - calculul mediei aritmetice simple a valorilor din fiecare celulă;
- b - calculul mediei aritmetice simple a valorilor medii din toate celulele rețelei.

Media estimată prin *declustering celular* depinde de dimensiunea celulelor rețelei rectangulare. Dacă celulele sunt prea mici, în fiecare celulă va fi un singur punct de observație și fiecare valoare va avea aceeași pondere. Dacă celulele sunt așa de mari încât toate valorile intră într-o singură celulă, din nou toate valorile vor avea aceeași pondere. Între aceste două situații extreme se află dimensiunile celulelor care permit evaluarea *mediei reprezentative*. Alegerea dimensiunilor optime ale celulei rețelei rectangulare se bazează pe analiza *hărții conturale a mediei globale* construită pentru diferite valori ale dimensiunilor celulei de discretizare a ariei explorate și în corelație cu *obiectivul estimării*.

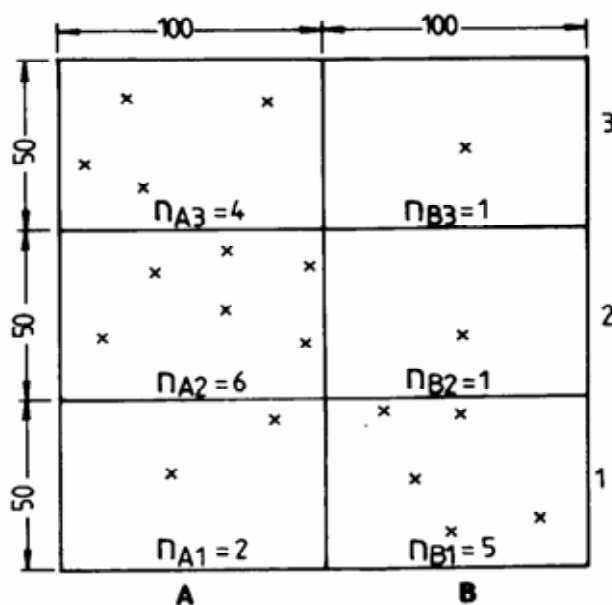


Fig. 2.5. Declustering celular

A20. Calculul valorii medii prin declustering celular

Să se calculeze valoarea medie a porozității unui strat explorat prin 121 de foraje amplasate într-o rețea neuniformă (**Fig.2.6**) din care au fost prelevate 484 de probe (câte 4 din fiecare foraj). Ținând seama că din strat va fi exploatat petrol, să se recomande porozitatea optimă pentru calculul rezervelor.

Rezolvare:

Cele 121 de foraje executate pot fi încadrate într-un dreptunghi cu lățimea de 30 km (pe direcția VE) și lungimea de 60 km (pe direcția NS).

Caracteristicile statistice elementare ale celor 484 de valori măsurate ale porozității totale sunt:

- distribuție cvasi-normală (**Fig.2.7**);
- amplitudinea selecției : $A = 40,7 - 17,1 = 23,6\%$
- coeficientul de asimetrie $\beta_1 = 0,15$
- media de selecție, la o eroare de genul I, $\alpha = 5\%$:

$$m = 30,7\% \pm 0,49$$

Media astfel calculată nu ține seama de distribuția neuniformă a forajelor de explorare. Pentru evaluarea influenței acestei distribuții neuniforme prin metoda declusteringului celular se construiesc 200 de rețele de discretizare cu *celule rectangulare* (având laturile multipli de 3 km) pornind de la cea mai mică - cu suprafața de 9 km² (3 km x 3 km) - până la cea mai mare - cu o suprafață de 1800 km² (30 km x 60 km).

Cele 200 de *valori medii ale porozității totale calculate* pentru cele 200 de rețele de discretizare au următoarele caracteristici statistice elementare:

- distribuție normală (**Fig.2.8**);
- amplitudinea de selecție:
 $A = 31,04 - 28,552 = 2,52\%$
- coeficientul de asimetrie $\beta_1 = 0,01$;
- media de selecție, la o eroare de genul I, $\alpha = 5\%$:

$$m = 29,5\% \pm 0,06$$

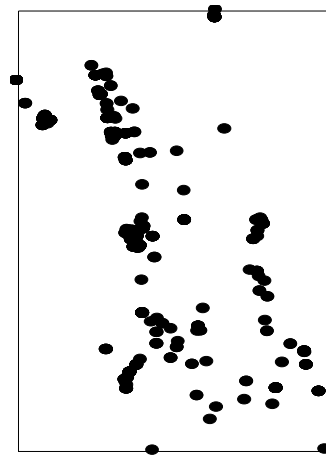


Fig.2.6. Amplasarea forajelor

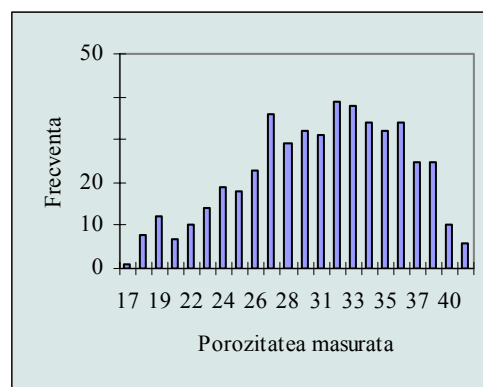


Fig.2.7. Histograma porozităților măsurate

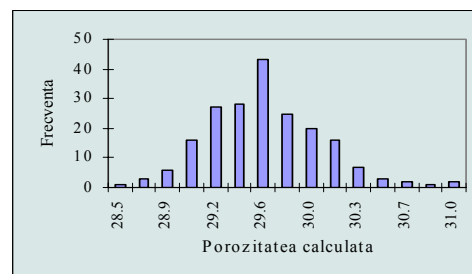


Fig.2.8. Histograma porozităților calculate prin declustering celular.

Referitor la influența dimensiunii celulelor de discretizare, trebuie remarcată

tendința de creștere a mediei porozității proporțional cu creșterea suprafeței celulelor de discretizare (**Fig.2.9**). Această corelație este de tip empiric și corespunde situației concrete a perimetrului explorat. Nu trebuie făcută nici o generalizare. Este posibil ca pentru un alt perimetru tendința să fie inversă sau să nu existe o corelație globală între dimensiunea celulei de discretizare și valoarea porozității medii calculate.

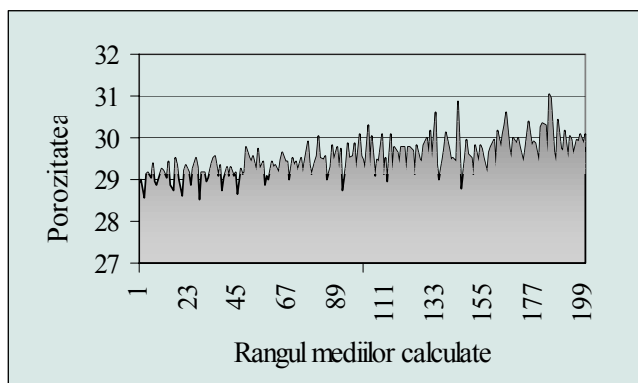


Fig.2.9. Corelația între porozitatea calculată și dimensiunea celulelor de discretizare.

Tot de natură empirică este și dependența valorii medii a porozității calculate de dimensiunea celulelor de discretizare și orientarea lor în spațiu, dependență reflectată în *harta cu izolinii* a valorii medii a porozității (**Fig.2.10**).

Harta conturală a porozității medii permite stabilirea cu rapiditate a valorii medii a porozității în

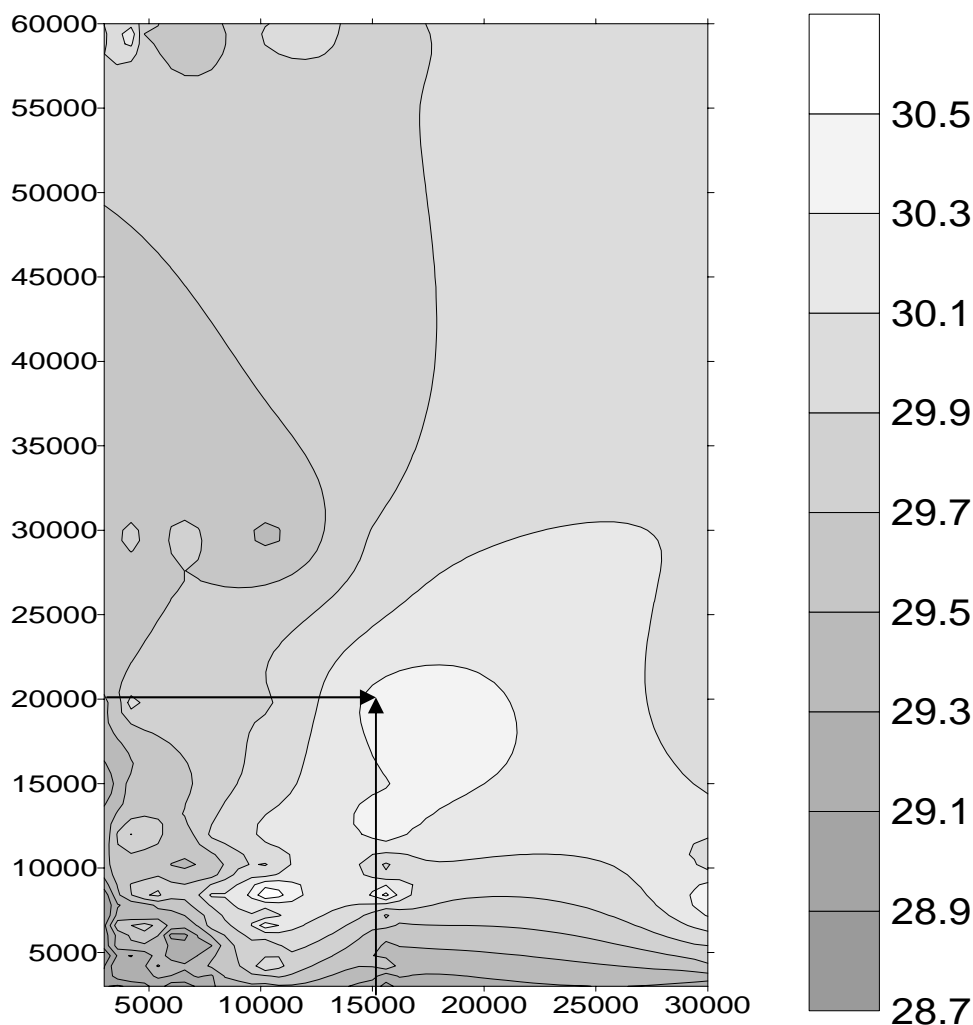


Fig.2.10. Harta conturală a porozității medii calculate prin declustering celular

funcție de dimensiunea celulei de discretizare.

Astfel, pentru discretizarea suprafeței prin celule rectangulare cu lățimea de 1500 m (pe direcția VE) și lungimea de 20000 m (pe direcția NS) se obține o valoare medie a porozității cuprinsă între 30,3% și 30,5% (conform scării grafice asociată hărții). Valoarea medie se citește de pe harta cu izolinii a mediei, la intersecția celor două săgeți corespunzătoare dimensiunilor celulei de discretizare (**Fig. 2.10**).

Recomandarea porozității optime pentru estimarea rezervelor de petrol se poate face prin minimizarea riscului de supraestimare a acestor rezerve. În această ipoteză de lucru se va recomanda *valoarea medie minimă a porozității calculate prin declustering celular*, adică 28,52%, cu 1% mai mică decât media de selecție (29,55%).

COMENTARIU

Estimarea globală este cel mai grosier nivel de estimare a distribuției spațiale.

Chiar dacă intenționăm să realizăm o estimare detaliată a unei structuri spațiale, nu este prudent să nu facem în primul rând o estimare globală. Evaluarea unei valori medii, valabilă pentru întreaga zonă, este un bun element de control al estimărilor detaliate pe care le realizăm ulterior. Abaterile exagerate ale valorilor locale estimate față de media generală sunt semnale pentru posibile erori determinate de erori de operare sau de neadecvare a instrumentelor de interpolare (ex.: variogramă, distanță fractală etc.).

Cele două metode prezentate elimină erorile introduse de distribuția neuniformă a punctelor de probare, fără a ține seama de efectul nestaționarității care determină “devierea” mediei datorită cumulării efectului distribuției neuniforme și a prezenței tendinței.

Imaginați-vă că în zona cercetată pentru distribuția spațială a conținutului de hidrocarburi se remarcă o tendință de creștere a acestuia de la sud spre nord!

Estimarea unei valori medii reprezentative pentru întreaga zonă este imposibilă în prezența nestaționarității:

- *dacă în zona sudică sunt mai multe puncte de probare decât în cea nordică media conținutului de hidrocarburi va fi subestimată;*
- *dacă în zona nordică sunt mai multe puncte de probare decât în zona sudică media este supraestimată;*
- *dacă densitatea punctelor de probare este uniformă pentru întreaga zonă, valoarea medie calculată este influențată de parametrii tendinței (pantă și direcție).*

*În astfel de situații, în etapa estimării globale se stabilește forma analitică a tendinței (vezi **A12**) careia i se asociază o eroare pe baza prelucrării reziduului. Dacă reziduul este un semnal aleator cu distribuție normală, atunci identificarea și separarea tendinței a fost corect realizată. În orice punct al zonei studiate valoarea cea mai probabilă se va estima cu ajutorul ecuației tendinței și a valorii medii a reziduului.*

2.2. ESTIMAREA PUNCTUALĂ

Kriging-ul este metoda topo-probabilistă care constă în găsirea celei mai bune estimări lineare posibile a valorii medii într-un punct pe baza valorilor disponibile din vecinătatea acestuia.

Kriging-ul realizează o *ponderare* a acestor valori în așa fel încât *varianța* de estimare rezultată să fie *minimă*, ținând seama de geometria punctelor de observație și de variabilitatea spațială. În mare, așa cum este natural, kriging-ul va atribui ponderi mari valorilor apropiate și ponderi mici valorilor depărtate. Această regulă intuitivă poate fi uneori mascată de *efectul de ecranare* și de *transferul de influență*.

Pentru a face posibilă estimarea prin kriging a ponderilor acordate valorilor măsurate este necesară acceptarea unor ipoteze asupra caracteristicilor variabilei studiate, sintetizate în funcția de covarianță sau variogramă a funcției aleatoare a cărei unică realizare disponibilă se presupune a fi eșantionul de date.

Caracteristica principală a kriging-ului nu este numai valoarea minimă a varianței de estimare care presupune utilizarea celei mai mari părți a informației disponibile, deci obținerea celei mai bune estimări, dar și caracterul nedeviat al acesteia.

Obiectivele kriging-ului sunt irealizabile fără apelarea la modelul funcției aleatoare, eroarea de estimare fiind nedeterminabilă datorită necunoașterii valorii reale a variabilei în punctul de estimare.

Deoarece *media erorilor* (m_r) și *varianța de estimare* (σ_r^2) sunt necunoscute, în kriging se operează cu *media erorilor* și *varianța de estimare* a modelului (\tilde{m}_r și $\tilde{\sigma}_r^2$).

Stabilirea ecuațiilor pe baza cărora se calculează ponderile w_i din formula (2.1) implică transpunerea în cadrul *modelului funcției aleatoare* a *erorii de estimare* și a *varianței erorii de estimare*.

Eroarea de estimare. Pentru fiecare punct în care nu dispunem de o valoare măsurată, prin kriging se estimează valoarea necunoscută utilizând o combinație lineară a valorilor cunoscute de forma (2.1).

Notăm cu r_i eroarea unei anumite estimări punctuale ($i=1,2,\dots,k$) și o definim ca diferența dintre valoarea estimată (v_i^*) și cea reală (v_i):

$$r_i = v_i^* - v_i, \quad (2.5)$$

Media tuturor erorilor de estimare punctuală este:

$$m_r = \frac{1}{k} \sum_{i=1}^k r_i = \frac{1}{k} \sum_{i=1}^k (v_i^* - v_i). \quad (2.6)$$

Utilizarea expresiei (2.6) pentru calcule nu este posibilă deoarece nu se cunosc valorile adevărate ale variabilei în punctele în care nu avem măsurători.

Soluționarea problemei se bazează pe apelarea la *modelul funcției aleatoare*: se consideră că atât valorile necunoscute cât și cele cunoscute sunt realizările unei funcții aleatoare staționare. În fiecare punct de observație avem amplasată o funcție aleatoare $V(p_i)$ și de asemenea câte una în fiecare punct de estimare $V(p_o)$.

Fiecare variabilă aleatoare are aceeași lege de probabilitate și în fiecare locație speranța matematică $E\{V\}$ este aceeași. Corelația între valorile fiecărei perechi de

variabile aleatoare depinde numai de distanța dintre ele. *Covarianța* unei perechi de variabile aleatoare separate prin distanța h o notăm cu $c(h)$.

Fiecare valoare măsurată este considerată ca o realizare a unei variabile aleatoare. Valorile estimate care sunt combinații lineare ale acestor valori sunt și ele variabile aleatoare:

$$V^*(p_0) = \sum_{i=1}^n w_i V(p_i) . \quad (2.7)$$

În mod similar, *erorile de estimare* definite ca diferență între *valoarea estimată* și cea *reală*, ambele *variabile aleatoare*, sunt și ele variabile aleatoare:

$$R(p_0) = V^*(p_0) - V(p_0) . \quad (2.8)$$

Substituind expresia (2.7) în expresia (2.8) se poate exprima eroarea de estimare prin intermediul variabilelor aleatoare originale:

$$R(p_0) = \sum_{i=1}^n w_i V(p_i) - V(p_0) . \quad (2.9)$$

Eroarea comisă la estimarea unei valori necunoscute în p_0 este deci o realizare a variabilei aleatoare $R(p_0)$ iar pentru ca estimarea în orice locație să fie nedeviată speranța matematică a erorii trebuie să fie zero:

$$E\{R(p_0)\} = E\left\{\sum_{i=1}^n w_i V(p_i) - V(p_0)\right\} = \sum_{i=1}^n w_i \cdot E\{V(p_i)\} - E\{V(p_0)\} = 0 \quad (2.10)$$

Deoarece este presupusă staționaritatea funcției aleatoare, pentru ca estimarea să fie nedeviată suma ponderilor w_i trebuie să fie unitară:

$$E\{R(p_0)\} = \sum_{i=1}^n w_i E\{V\} - E\{V\} = E\{V\} \left(\sum_{i=1}^n w_i - 1\right) = 0 \quad (2.11)$$

$$\sum_{i=1}^n w_i = 1$$

Varianța erorii de estimare. Varianța erorii de estimare pentru un set de estimări poate fi scrisă sub forma:

$$\sigma_R^2 = \frac{1}{k} \sum_{i=1}^k (r_i - m_r)^2 = \frac{1}{k} \sum_{i=1}^k \left[v_i^* - v_i - \frac{1}{k} \sum_{i=1}^k (v_i^* - v_i) \right]^2 . \quad (2.12)$$

Dacă presupunem și caracterul nedeviat al estimării (2.11), rezultă că:

$$\sigma_R^2 = \frac{1}{k} \sum_{i=1}^k [v_i^* - v_i]^2 . \quad (2.13)$$

Expresia (2.13) nu este operațională deoarece nu se cunosc valorile reale în punctele de estimare (v_i). Pentru a rezolva problema se apelează din nou la *modelul funcției aleatoare*.

Se pornește de la $(n+1)$ variabile aleatoare, n din ele modelând comportarea fenomenului în locațiile cunoscute și a $(n+1)$ -a în punctul p_o unde se realizează estimăția. Estimatorul $V^*(p_o)$ este tot o variabilă aleatoare deoarece este o combinație lineară de variabile aleatoare:

$$V^*(p_o) = \sum_{i=1}^k w_i \cdot V(p_i) . \quad (2.14)$$

Diferența dintre valoarea reală și cea estimată este de asemenea o variabilă aleatoare:

$$R(p_o) = V^*(p_o) - V(p_o) , \quad (2.15)$$

care poate fi dezvoltată sub forma:

$$Var\{R(p_o)\} = Cov\{V^*(p_o) \cdot V^*(p_o)\} - 2Cov\{V^*(p_o) \cdot V(p_o)\} + Cov\{V(p_o) \cdot V(p_o)\} \quad (2.16)$$

Evaluarea rezidului este posibilă doar în cadrul modelului funcției aleatoare staționare. Acest model permite evaluarea covarianței sau variogramei pentru valorile necunoscute din punctele de estimare pe baza modelului de covarianță sau variogramă dedus din valorile măsurate.

Primul termen din relația (2.16) reprezintă covarianța valorii estimate a variabilei studiate cu ea însăși, adică varianța valorii estimate, ea însăși o combinație lineară de variabile aleatoare:

$$Cov\{V^*(p_o) \cdot V^*(p_o)\} = Var\left\{\sum_{i=1}^n w_i V(p_i)\right\} = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \tilde{c}_{ij} , \quad (2.17)$$

în care:

- \tilde{c}_{ij} - covarianța modelată dintre două puncte p_i și p_j situate la distanța h_{ij} și în care se cunosc valorile variabilei;
- w_i și w_j - ponderile acordate valorilor măsurate în punctele p_i și p_j .

Al doilea termen din relația (2.16) poate fi descompus sub forma:

$$\begin{aligned} 2Cov\{V^*(p_o) \cdot V(p_o)\} &= 2Cov\left\{\left(\sum_{i=1}^n w_i \cdot V(p_i)\right) \cdot V(p_o)\right\} = \\ &= 2E\left\{\sum_{i=1}^n w_i \cdot V(p_i) \cdot V(p_o)\right\} - 2E\left\{\sum_{i=1}^n w_i \cdot V(p_i)\right\} \cdot E\{V(p_o)\} = \\ &= 2\sum_{i=1}^n w_i E\{V(p_i) \cdot V(p_o)\} - 2w_i E\{V(p_i) \cdot V(p_o)\} = \\ &= 2\sum_{i=1}^n w_i (E\{V(p_i) \cdot V(p_o)\} - E\{V(p_i)\} \cdot E\{V(p_o)\}) = \\ &= 2\sum_{i=1}^n w_i Cov\{V(p_i) \cdot V(p_o)\} = 2\sum_{i=1}^n w_i \cdot \tilde{c}_{i0} . \end{aligned} \quad (2.18)$$

Al treilea termen din ecuația (2.16), prin analogie cu primul este varianța valorii reale din punctul de estimare p_o care se exprimă prin intermediul modelului de covarianță sub forma:

$$\text{Cov}\{V(p_o) \cdot V(p_o)\} = \tilde{\sigma}^2 \quad (2.19)$$

Pe baza ecuațiilor (2.17), (2.18) și (2.19) se obține expresia varianței erorilor de estimare care permite evaluarea ei pe baza modelului de covarianță :

$$\tilde{\sigma}_R^2 = \tilde{\sigma}^2 + \sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{c}_{ij} - 2 \sum_{i=1}^n w_i \cdot \tilde{c}_{i0} , \quad (2.20)$$

în care:

\tilde{c}_{i0} - covarianța modelată între punctele p_i în care se cunosc valorile variabilei și p_o în care se estimează valoarea variabilei, situate la distanța h_{i0} .

Realizarea obiectivului operațional principal al estimării punctuale (evaluarea ponderilor w_i cu care valorile măsurate (v_i) participă la estimarea variabilei studiate), în funcție de caracteristica structurii spațiale a variabilei studiate, se poate face prin:

- *kriging punctual ordinar*, dacă variabila studiată este *staționară*, cu repartiție normală;
- *kriging punctual universal*, dacă variabila studiată este *nestaționară*, cu repartiție normală.

2.2.1. KRIGING PUNCTUAL ORDINAR

Calculul ponderilor din combinația liniară (2.1) care asigură estimarea nedeviată și minimizarea erorii de estimare poate fi realizat prin *kriging punctual ordinar (staționar)* utilizând toate cele trei funcții de continuitate: *covarianță*, *variogramă* și *corelogramă*.

Ecuatiile sistemului pentru kriging-ul ordinar în funcție de covarianță.

Calculul ponderilor acordate valorilor măsurate (w_i) în estimarea prin *kriging ordinar* presupune asigurarea simultană a estimării nedeviate:

$$m_r = 0 \quad (2.21)$$

și a minimului varianței erorii de estimare:

$$\sigma_R^2 - \text{minimum} \quad (2.22)$$

Minimizarea varianței erorii de estimare nu se poate realiza prin simpla anulare a derivatelor parțiale în raport cu ponderile w_i , deoarece trebuie asigurată și respectarea condiției din ecuația (2.21). Asigurarea condiționării suplimentare din ecuația (2.22) este realizată prin utilizarea *parametrului lui Lagrange* care convertește problema minimizării condiționate într-o problemă de minimizare fără condiții.

Minimizarea varianței de estimare din expresia (2.20) cu condiția de estimare nedeviată din ecuația (2.21) conduce prin egalarea cu zero a derivatelor parțiale în raport cu necunoscutele (w_i) la un sistem nedeterminat de $(n+1)$ ecuații cu n necunoscute. Pentru soluționarea problemei se introduce o nouă necunoscută în ecuația (2.22) numită parametrul lui Lagrange (μ):

$$\tilde{\sigma}_R^2 = \tilde{\sigma}^2 + \sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{c}_{ij} - 2 \sum_{i=1}^n w_i \cdot \tilde{c}_{i0} + 2\mu \left(\sum_{i=1}^n w_i - 1 \right). \quad (2.23)$$

Cantitatea adăugată, al patrulea termen al ecuației, nu modifică ecuația, el fiind nul prin chiar condiția de estimare nedeviată care se adaugă în acest mod minimizării varianței erorii de estimare.

Ecuția varianței erorii de estimare în forma (2.23) este o ecuație cu $(n+1)$ necunoscute a cărei minimizare se realizează prin anularea celor $(n+1)$ derivate parțiale în raport cu $w_1, w_2, \dots, w_n, \mu$ ecuații ce constituie sistemul de kriging ordinar:

$$\begin{cases} \frac{\partial(\tilde{\sigma}_R^2)}{\partial w_1} = 0 \\ \vdots \\ \frac{\partial(\tilde{\sigma}_R^2)}{\partial w_n} = 0 \\ \frac{\partial(\tilde{\sigma}_R^2)}{\partial \mu} = 0 \end{cases} \quad (2.24)$$

Calculul derivatei în raport cu w_1 , desfășurat separat pentru cei patru termeni ai varianței erorii de estimare dată de ecuația (2.23), conduce la următoarele rezultate:

- primul termen:

$$\frac{\partial(\tilde{\sigma}^2)}{\partial w_1} = 0 ; \quad (2.25)$$

- al doilea termen:

$$\begin{aligned} \frac{\partial \left(\sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{c}_{ij} \right)}{\partial w_1} &= \frac{\partial \left(w_1^2 \tilde{c}_{11} + 2w_1 \sum_{j=2}^n w_j \cdot \tilde{c}_{1j} \right)}{\partial w_1} = ; \\ &= 2w_1 \tilde{c}_{11} + 2 \sum_{j=2}^n w_j \cdot \tilde{c}_{1j} = 2 \sum_{j=1}^n w_j \cdot \tilde{c}_{1j} \end{aligned} \quad (2.26)$$

- al treilea termen:

$$\frac{\partial \left(2 \sum_{i=1}^n w_i \cdot \tilde{c}_{i0} \right)}{\partial w_1} = 2 \tilde{c}_{10} ; \quad (2.27)$$

- al patrulea termen:

$$\frac{\partial \left(2\mu \left(\sum_{i=1}^n w_i - 1 \right) \right)}{\partial w_1} = 2\mu . \quad (2.28)$$

Prin combinarea ecuațiilor (2.26 - 2.28) se obține expresia pentru derivata în raport cu w_1 :

$$\frac{\partial(\tilde{\sigma}_R^2)}{\partial w_1} = 2 \sum_{j=1}^n w_j \cdot \tilde{c}_{ij} - 2\tilde{c}_{i0} + 2\mu . \quad (2.29)$$

Primele n derivate în raport cu w_i ($i = 1, 2, \dots, n$) au forma generală:

$$\sum_{j=1}^n w_j \cdot \tilde{c}_{ij} + \mu = \tilde{c}_{i0}, \forall i = 1, 2, \dots, n . \quad (2.30)$$

Derivata în raport cu parametrul lui Lagrange, a $(n+1)$ -a derivată, are forma:

$$\frac{\partial(\tilde{\sigma}_R^2)}{\partial \mu} = \frac{\partial(2\mu(\sum_{i=1}^n w_i - 1))}{\partial \mu} = 2(\sum_{i=1}^n w_i - 1) . \quad (2.31)$$

Sistemul de kriging ordinar se obține prin anularea celor $(n+1)$ derivate parțiale date de ecuațiile (2.30) și (2.31) și are forma:

$$\begin{cases} \frac{\partial(\tilde{\sigma}_R^2)}{\partial w_1} = 2 \sum_{i=1}^n w_i \cdot \tilde{c}_{1i} - 2\tilde{c}_{10} + 2\mu = 0 \\ \vdots \\ \frac{\partial(\tilde{\sigma}_R^2)}{\partial w_n} = 2 \sum_{i=1}^n w_i \cdot \tilde{c}_{ni} - 2\tilde{c}_{n0} + 2\mu = 0 \\ \frac{\partial(\tilde{\sigma}_R^2)}{\partial \mu} = 2 \sum_{i=1}^n w_i - 1 = 0 \end{cases} \quad (2.32)$$

Prin separarea coeficienților cunoscuți și a necunoscutelor, sistemul de kriging poate fi scris sub forma matricială:

$$\begin{bmatrix} \tilde{c}_{11} & \tilde{c}_{12} & \cdots & \tilde{c}_{1n} & 1 \\ \tilde{c}_{21} & \tilde{c}_{22} & \cdots & \tilde{c}_{2n} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \tilde{c}_{n1} & \tilde{c}_{n2} & \cdots & \tilde{c}_{nn} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} \tilde{c}_{10} \\ \tilde{c}_{20} \\ \vdots \\ \tilde{c}_{n0} \\ 1 \end{bmatrix} . \quad (2.33)$$

Dacă notăm cu C matricea coeficienților, cu W vectorul necunoscutelor și cu D vectorul termenilor liberi, sistemul (2.33) poate fi scris sub forma:

$$C \cdot W = D \quad (2.34)$$

a cărei soluție este:

$$W = C^{-1} \cdot D \quad (2.35)$$

Varianța erorii de estimare a variabilei în punctul p_o este mai mică decât varianța dispersiei totale a funcției aleatoare (σ^2), acest lucru fiind determinat de existența punctelor p_i în care cunoaștem valorile acesteia.

Calculul valorii minime a varianței erorii de estimare poate utiliza relația (2.23), dar pentru a găsi o expresie în raport numai cu valorile măsurate se pleacă de la ecuația (2.30) în care ambii membri se multiplică cu w_i :

$$w_i \left(\sum_{j=1}^n w_j \cdot \tilde{c}_{ij} + \mu \right) = w_i \cdot \tilde{c}_{i0} \quad (2.36)$$

și se însumează pentru toate cele n puncte de observație, rezultând:

$$\sum_{i=1}^n w_i \sum_{j=1}^n w_j \cdot \tilde{c}_{ij} + \sum_{i=1}^n w_i \cdot \mu = \sum_{i=1}^n w_i \cdot \tilde{c}_{i0} \quad (2.37)$$

care sub forma:

$$\sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{c}_{ij} = \sum_{i=1}^n w_i \cdot \tilde{c}_{i0} + \mu \quad (2.38)$$

se înlocuiește în (2.23) și se obține:

$$\tilde{\sigma}_R^2 = \tilde{\sigma}^2 - \left(\sum_{i=1}^n w_i \cdot \tilde{c}_{i0} + \mu \right), \quad (2.39)$$

care în formă matricială este:

$$\tilde{\sigma}_R^2 = \tilde{\sigma}^2 - w \cdot D \quad (2.40)$$

Ecuatiile sistemului de kriging în funcție de variogramă. În aceleași ipoteze care au permis deducerea expresiei varianței erorii de estimare (2.23) poate fi utilizată și *variograma* a cărei relație de definiție este:

$$\gamma_{ij} = \frac{1}{2} E \{ [V(p_i) - V(p_j)]^2 \}, \quad (2.41)$$

care pentru evaluarea varianței erorii de estimare poate fi scrisă sub forma:

$$\begin{aligned} \gamma_{ij} &= \frac{1}{2} E \{ [(V(p_i) - V(p_o)) - (V(p_j) - V(p_o))]^2 \} = \\ &= \frac{1}{2} E \{ [V(p_i) - V(p_o)]^2 \} + \frac{1}{2} E \{ [V(p_j) - V(p_o)]^2 \} - \\ &- E \{ [V(p_i) - V(p_o)][V(p_j) - V(p_o)] \} = \\ &= \tilde{\gamma}_{i0} + \tilde{\gamma}_{j0} - E \{ [V(p_i) - V(p_o)][V(p_j) - V(p_o)] \} \end{aligned} \quad (2.42)$$

Varianța erorii de estimare în această variantă este dată de relația:

$$\begin{aligned}\tilde{\sigma}_R^2 &= E\left\{\left[V^*(p_0) - V(p_0)\right]^2\right\} = E\left\{\left[\sum_{i=1}^n w_i \cdot V(p_i) - \sum_{i=1}^n w_i \cdot V(p_0)\right]^2\right\} = \\ &= \sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot E\left\{\left[V(p_i) - V(p_0)\right]\left[V(p_j) - V(p_0)\right]\right\}\end{aligned}\quad (2.43)$$

în care, utilizând forma variogramei din ecuația (2.42), se obține:

$$\begin{aligned}\tilde{\sigma}_R^2 &= E\left\{\left[V^*(p_0) - V(p_0)\right]^2\right\} = -\sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{\gamma}_{ij} + \\ &+ \sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{\gamma}_{i0} + \sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{\gamma}_{j0} = \\ &= -\sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{\gamma}_{ij} + \sum_{j=1}^n w_j \left(\sum_{i=1}^n w_i \cdot \tilde{\gamma}_{i0} + \sum_{i=1}^n w_i \cdot \tilde{\gamma}_{j0}\right) = \\ &= -\sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{\gamma}_{ij} + 2\sum_{i=1}^n w_i \cdot \tilde{\gamma}_{i0}\end{aligned}\quad (2.44)$$

Pentru minimizarea varianței erorii de estimare, în condiția de estimare nedeviată (2.21), utilizând parametrul lui Lagrange (μ), relația corespunzătoare ecuației (2.23) scrisă pentru variogramă este:

$$\tilde{\sigma}_R^2 = -\frac{1}{2} \left(\sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{\gamma}_{ij} + 2\sum_{i=1}^n w_i \cdot \tilde{\gamma}_{i0} \right) - \mu \left(\sum_{i=1}^n w_i - 1 \right) \quad (2.45)$$

Minimizarea varianței erorii de estimare scrisă sub forma (2.45) se realizează prin anularea derivatelor în raport cu cele $(n+1)$ necunoscute: $w_1, w_2, \dots, w_n, \mu$.

În mod analog cu ecuațiile (2.26-2.28) se obțin derivatele în raport cu ponderile w sub forma:

$$\frac{\partial(\tilde{\sigma}_R^2)}{\partial w_i} = \frac{1}{2} \left(-2\sum_{j=1}^n w_j \cdot \tilde{\gamma}_{ij} + 2\tilde{\gamma}_{i0} \right) - \mu, \forall i = 1, 2, \dots, n \quad (2.46)$$

iar pentru derivata în raport cu μ

$$\frac{\partial(\tilde{\sigma}_R^2)}{\partial \mu} = -\sum_{i=1}^n w_i + 1 \quad (2.47)$$

Sistemul de kriging în raport cu variograma se obține prin anularea celor $(n+1)$ derivate parțiale din ecuațiile (2.46) și (2.47) și are forma:

$$\begin{cases} \sum_{j=1}^n w_j \cdot \tilde{\gamma}_{ij} + \mu = \tilde{\gamma}_{i0}, \forall i = 1, 2, \dots, n \\ \sum_{i=1}^n w_i = 1 \end{cases} \quad (2.48)$$

care sub formă matricială poate fi scris:

$$\begin{bmatrix} \tilde{\gamma}_{11} & \tilde{\gamma}_{12} & \cdots & \tilde{\gamma}_{1n} & 1 \\ \tilde{\gamma}_{21} & \tilde{\gamma}_{22} & \cdots & \tilde{\gamma}_{2n} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \tilde{\gamma}_{n1} & \tilde{\gamma}_{n2} & \cdots & \tilde{\gamma}_{nn} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} \tilde{\gamma}_{10} \\ \tilde{\gamma}_{20} \\ \vdots \\ \tilde{\gamma}_{n0} \\ 1 \end{bmatrix} \quad (2.49)$$

Valoarea minimă a *varianței erorii de estimare* în raport cu valorile măsurate este dată de relația:

$$\tilde{\sigma}_R^2 = \sum_{i=1}^n w_i \cdot \tilde{\gamma}_{i0} + \mu \quad (2.50)$$

dedusă în mod analog cu relația (2.40).

Ecuatiile sistemului de kriging în funcție de corelogramă. Între corelogramă și covarianță există relația:

$$\tilde{\rho}_{ij} = \frac{\tilde{c}_{ij}}{\tilde{\sigma}^2}, \quad (2.51)$$

care este valabilă pentru un model de funcție aleatoare în care toate variabilele aleatoare au aceeași medie și aceeași dispersie. Valabilitatea acestei relații ne permite să scriem sistemul de kriging și în raport cu corelograma.

Dedus în mod analog cu sistemele (2.33) și (2.49) sistemul poate fi scris:

$$\begin{cases} \sum_{j=1}^n w_j \cdot \tilde{\rho}_{ij} + \mu = \tilde{\rho}_{i0}, \forall i = 1, 2, \dots, n \\ \sum_{i=1}^n w_i = 1 \end{cases} \quad (2.52)$$

sau sub formă matricială:

$$\begin{bmatrix} \tilde{\rho}_{11} & \tilde{\rho}_{12} & \cdots & \tilde{\rho}_{1n} & 1 \\ \tilde{\rho}_{21} & \tilde{\rho}_{22} & \cdots & \tilde{\rho}_{2n} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \tilde{\rho}_{n1} & \tilde{\rho}_{n2} & \cdots & \tilde{\rho}_{nn} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} \tilde{\rho}_{10} \\ \tilde{\rho}_{20} \\ \vdots \\ \tilde{\rho}_{n0} \\ 1 \end{bmatrix}. \quad (2.53)$$

În cazul utilizării corelogramei *varianța minimă a erorii de estimare* se calculează cu relația:

$$\tilde{\sigma}_R^2 = \sum_{i=1}^n w_i \cdot \tilde{\rho}_{i0} + \mu \quad (2.54)$$

În practică, datorită flexibilității variogramei se preferă utilizarea acesteia pentru estimarea *ponderilor* (w_i) (sistemul 2.49) și *a varianței minime de estimare* ($\tilde{\sigma}_k^2$) (ecuația 2.50).

Componentele sistemelor de kriging. Înțelegerea rolului matricilor C și D ale sistemelor de kriging (ec. 2.34) este determinantă în alegerea corectă a modelelor de covarianță sau variogramă, indispensabile estimărilor corecte.

Pentru mulți practicieni dezvoltările matematice necesare stabilirii sistemelor de kriging par complicate și nu reușesc să aducă o clarificare a rolului matricilor C și D , motiv pentru care în continuare vom încerca să dăm o explicație care să permită formarea unei imagini intuitive a acestui rol. Existența acestei înțelegeri intuitive a rolului matricilor C și D permite practicienilor realizarea unor ajustări care în ciuda aparentei lipse de rigoare teoretică pot să îmbunătățească metoda de estimare.

Matricea C înregistrează *distanțele statistice* dintre toate punctele de probare furnizând sistemului de kriging informații în legătură cu gruparea spațială a punctelor de probare.

Dacă matricea de kriging (C) este construită cu ajutorul covarianțelor (2.33) pentru două puncte de probare apropiate, valoarea în matrice va fi mare, iar pentru puncte îndepărtate valoarea va fi mică. În acest mod matricea C permite eliminarea influenței distanței neuniforme dintre punctele de probare, asupra estimărilor. Prin intermediul covarianțelor se realizează o ponderare similară cu *metoda declustering-ului poligonal sau celular*. Spre deosebire de metodele de declustering, metoda covarianțelor realizează o ponderare bazată pe structura intrinsecă a variabilei și nu pe o regulă arbitrară aleasă indiferent de variabila studiată.

Matricea D produce o ponderare a valorilor măsurate similară cu cea produsă de *metoda inversului distanței*. Covarianțele din matricea D descresc proporțional cu creșterea distanței dintre punctul în care se face estimarea (p_o) și punctele de observație (p_i ; $i = 1, 2, \dots, n$). Spre deosebire de ponderarea din cadrul metodei inversului distanței, în matricea D covarianțele se calculează în raport cu o *distanță statistică*, funcția de covarianță fiind valabilă în medie pentru tot domeniul spațial în care au fost măsurate valorile disponibile ale variabilei studiate.

Multiplicarea $C^{-1} \cdot D$ ajustează ponderile din D prin eliminarea redundanțelor dintre punctele de măsurare determinate de distribuția lor neuniformă și de tipul de continuitate spațială cuantificată prin intermediul modelului de covarianță, variogramă sau corelogramă.

Rezultă că sistemul de kriging ia în considerare simultan: *gruparea spațială* a punctelor, decodificată prin intermediul *distanței statistice*, și *tipul de continuitate* relevat de setul de date disponibil și concretizat în *modelul de covarianță*.

Kriging cu date incerte. Pentru estimarea prin kriging s-a presupus că toate valorile $v(p_i)$ ($i = 1, 2, \dots, n$) ale variabilei studiate sunt cunoscute fără nici o incertitudine. În realitate acest lucru este foarte rar și de cele mai multe ori datele provin din surse diferite sau sunt determinate prin metode diferite.

Analiza distribuției parametrilor hidrogeologici ai unui acvifer este obligată deseori să utilizeze valori ale conductivității hidraulice determinate prin pompări sau în laborator, ale grosimii acviferului determinate prin carotaj geofizic sau din carotajele mecanice. Studiile hidrochimice apelează la analize chimice realizate în laboratoare diferite sau prin metode diferite. Sunt doar câteva exemple în care gradul de încredere în valorile utilizate este diferit.

Kriging-ul poate utiliza astfel de date și ține cont de erorile asociate fiecărei valori (ε_i) dacă :

-erorile nu sunt sistematice:

$$E\{\varepsilon_i\} = 0 \quad i = 1, 2, \dots, n \quad (2.55)$$

-erorile nu sunt corelate între ele:

$$Cov\{\varepsilon_i, \varepsilon_j\} = 0, \forall i \neq j \quad (2.56)$$

-erorile nu sunt corelate cu mărimea valorilor măsurate:

$$Cov\{\varepsilon_i, V(p_i)\} = 0, \forall i, p_i \quad (2.57)$$

-erorile au o dispersie cunoscută σ_i .

Singura modificare în raport cu sistemele de kriging este aceea că ecuațiile vor conține pe diagonala principală un termen suplimentar $-\tilde{\sigma}_i^2$. Pentru *kriging-ul punctual ordinar în cazul datelor incerte* sistemul va avea forma:

$$\begin{cases} \sum_{j=1}^n w_i \cdot \tilde{\gamma}_{ij} - w_i \cdot \tilde{\sigma}_i^2 + \mu = \tilde{\gamma}_{i0}, i = 1, 2, \dots, n \\ \sum_{i=1}^n w_i = 1 \end{cases} \quad (2.58)$$

sau matricial:

$$\begin{bmatrix} \tilde{\gamma}_{11} - \tilde{\sigma}_1^2 & \tilde{\gamma}_{12} & \dots & \tilde{\gamma}_{1n} & 1 \\ \tilde{\gamma}_{21} & \tilde{\gamma}_{22} - \tilde{\sigma}_2^2 & \dots & \tilde{\gamma}_{2n} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \tilde{\gamma}_{n1} & \tilde{\gamma}_{n2} & \dots & \tilde{\gamma}_{nn} - \tilde{\sigma}_n^2 & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} \tilde{\gamma}_{10} \\ \tilde{\gamma}_{20} \\ \vdots \\ \tilde{\gamma}_{n0} \\ 1 \end{bmatrix} \quad (2.59)$$

Este evident că se pot utiliza simultan *date certe* și *incerte*, pentru *datele certe dispersia fiind nulă*, iar pentru cele *incerte diferită de zero*.

Erorile de măsurare pot afecta în mod uniform toate valorile disponibile, dispersia acestor erori fiind cuantificată de efectul de pepită al variogramei experimentale.

Utilizarea modelelor de variogramă cu efect de pepită ține seama de erorile de măsură fără să introducă nici o modificare în forma generală a ecuațiilor sistemului de kriging. În mod particular, un sistem de kriging pentru care efectul de pepită al modelului de variogramă utilizat este nul (datele fiind considerate certe) are pe diagonala principală a matricii C din ecuația (2.34) valori nule. Când efectul de pepită este diferit de zero, valorile lui apar pe diagonala principală a matricii C a sistemului de kriging. De reținut că atunci când punctul în care se face estimarea este un punct de observație efectul de pepită este zero.

Selectarea valorilor utilizate. Strategia de selectare a valorilor care sunt incluse în procedura de estimare a variabilei v într-un punct oarecare p_o trebuie să răspundă la minimum patru întrebări:

- dacă în vecinătatea punctului p_o sunt suficiente puncte de observație;
- dacă în vecinătatea punctului p_o sunt prea multe puncte de observație;
- dacă în vecinătatea punctului p_o sunt valori redundante;

- dacă în vecinătatea punctului p_o sunt valori relevante.

Primele trei sunt importante în mod deosebit pentru metodele de estimare de tipul kriging-ului, care pot utiliza un număr nelimitat de valori. Pentru limitarea efortului de calcul fără o diminuare a preciziei estimării, strategia obișnuită este de a utiliza toate valorile plasate într-o anumită *ferastră de selecție*.

Fereastra de selecție. Forma ferestrei de selecție este în general o elipsă centrată pe punctul p_o . Orientarea elipsei este determinată de direcțiile de anizotropie, semiaxa mare fiind paralelă cu direcția de continuitate maximă (**Fig.2.11a**). Dacă anizotropia nu este evidentă, elipsa se transformă în cerc, orientarea axelor este nerelevantă (**Fig.2.11b**).

Dimensiunea minimă a ferestrei de selecție se alege în așa fel încât să cuprindă suficiente valori pentru a permite o estimare satisfăcătoare și depinde în mod evident de geometria punctelor de observație.

Dacă punctele sunt amplasate într-o rețea rectangulară se poate calcula în funcție de parametrii rețelei cât de mare trebuie să fie elipsa ca să includă cel puțin patru puncte. Pentru o rețea neregulată semiaxa mare a elipsei (R) trebuie să fie mai mare decât distanța medie dintre punctele de probare, valoarea ei putând să fie aproximată cu formula:

$$R = ((\text{Aria totală a sprafeței probate})/(\text{Numărul punctelor de probare}))^{0,5} \quad (2.60)$$

Dimensiunea maximă a elipsei de căutare este determinată de dimensiunea matricii C și de domeniul spațial pe care poate fi considerată satisfăcătoare staționaritatea variabilei studiate.

Reducerea volumului de calcule care este proporțional cu *cubul numărului de valori individuale*, se poate realiza, fără reducerea numărului de puncte utilizate, prin combinarea valorilor brute după diferite scheme.

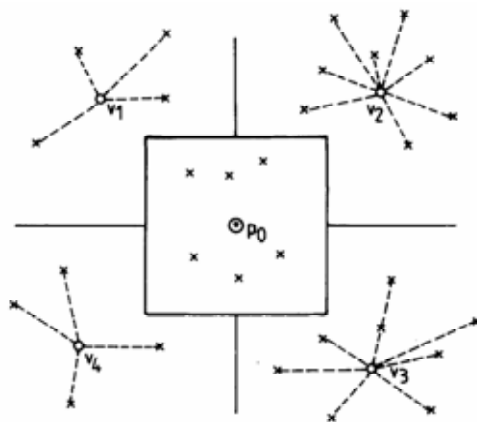


Fig. 2.12. Combinarea valorilor brute pentru reducerea volumului de calcule

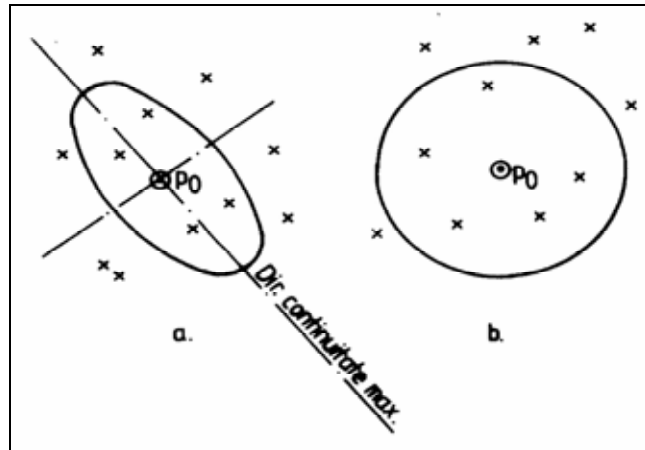


Fig.2.11. Forma ferestrei de estimare

În **Fig. 2.12** este prezentată o metodă de combinare a valorilor brute. Punctul p_o (în care se face estimarea) este încadrat într-un pătrat central care să cuprindă un număr rezonabil de puncte ce vor fi considerate individual în calcul, iar cele din sectoarele deschise se cumulează în patru valori compuse v_1, v_2, v_3, v_4 . Ponderea acordată valorii compuse poate fi distribuită în mod egal valorilor din care aceasta este compusă.

Modelarea variabilității prin intermediul funcțiilor aleatoare impune limitarea dimensiunii ferestrei de selecție pe domenii în care comportarea variabilei

poate fi considerată staționară. Simultan cu restrângerea ferestrei de selecție, staționaritatea devine mai plauzibilă iar diferența dintre proprietățile statistice ale selecției și cele ale modelului mai mică.

O concepție greșită, frecvent întâlnită în selectarea punctelor de calcul, este limitarea semiaxei mari a elipsei la raza variogramei. Experiența arată că dacă sunt puține valori în zona razei de influență a variogramei, utilizarea valorilor suplimentare aflate în afara acesteia deseori îmbunătățește precizia de estimare.

Redundanța valorilor selectate.
Eliminarea punctelor redundante din interiorul ferestrei de selecție se face în mod obișnuit prin împărțirea acesteia în 4, 8 sau 16 sectoare și limitarea numărului de puncte utilizate din fiecare sector. În acest mod se reduce efectul grupării neuniforme a punctelor de probare.

Numărul de sectoare în care se împarte fereastra de selecție este determinat de densitatea punctelor de observație și este evident că la o densitate mică a punctelor de observație se alege un număr redus de sectoare.

Cazul prezentat (**Fig.2.13**) este al unei ferestrei de selecție rectangulare separată în patru sectoare, din fiecare sector selectându-se două puncte. Selecția punctelor din fiecare sector s-a operat în cazul prezentat pe baza distanței dintre punctul de observație și cel de estimare (p_0). Sunt desenate toate punctele din fereastra de selecție și prin cercuri sunt marcate cele două puncte din fiecare sector care sunt selectate (**Fig.2.13a**) și separat punctele reținute pentru calcul (**Fig.2.13b**).

Problema redundanței punctelor selectate este soluționată în mod optim pentru kriging prin intermediul matricii C , dacă modelul de continuitate (covarianța, variograma sau corelograma) este bine ales. Efectul aplicării selecției sectorizate în acest caz este nul.

Tehnica sectorizării ferestrei de selecție este recomandată pentru metodele de estimare care nu utilizează matrici *declusterizante* în eliminarea efectului grupării punctelor de observație (ex.: metoda inversului distanței) sau în cazul kriging-ului, când pe baza datelor disponibile modelul de continuitate nu poate fi bine precizat.

Relevanța valorilor selectate.
Valorile utilizate în estimările punctuale sunt relevante dacă ele aparțin aceleiași populații statistice în care este încadrat și punctul estimat. Din nefericire, chiar dacă separarea populațiilor statistice se face minuțios utilizând analiza dispersională sau analiza discriminant (D.Scrădeanu, 1995), acest lucru nu poate fi verificat. De cele mai multe ori sunt necesare și decizii subiective care beneficiază de informații cu caracter calitativ iar obiectivul studiului contribuie și el la

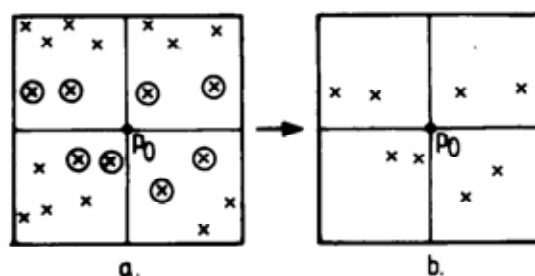


Fig. 2.13. Selectarea sectorizată a valorilor



Fig. 2.14. Selectarea valorilor relevante

separarea populațiilor.

Fig.2.14 prezintă o rețea de monitoring pentru calitatea apelor acviferului freatic din terasa unui curs de apă. În punctul p_o se intenționează estimarea conținutului de azotați proveniți din utilizarea îngrășămintelor chimice.

Ținând seama de combinatul de îngrășămintele chimice plasat la nord de punctul p_o și de direcțiile de curgere din acvifer, punctele 9, 10 și 11 din imediata vecinătate a acestuia vor trebui excluse, ele nefiind relevante pentru conținutul de azotați proveniți din utilizarea îngrășămintelor. Este evident că relevanța valorilor din punctele 9, 10 și 11 este maximă dacă se urmărește evaluarea în punctul p_o a influenței combinatului chimic asupra calității apei acviferului.

Alegerea corectă a punctelor relevante pentru estimare poate fi mai importantă decât alegerea metodei de estimare. Definirea domeniilor spațiale în conexiune cu relevanța valorilor este obligatoriu să fie primul pas în realizarea oricărei estimări spațiale.

Practica frecventă a utilizării aceleiași metode de selectare a valorilor pentru întreaga suprafață studiată nu este întotdeauna cea mai bună. Ceea ce este corect într-o anumită zonă poate fi incorect pentru alta iar adaptarea tehnicilor de selecție configurației particulare a structurilor studiate solicită programe automate cu un grad sporit de interactivitate.

Alegerea metodelor de estimare cât și selectarea punctelor utilizate în estimare trebuie să se bazeze pe o analiză detaliată a datelor disponibile. Ignorarea relevanței valorilor pentru obiectivele estimării și a redundanței introduse de distribuția lor neregulată afectează în mod diferențiat precizia estimării și pierderea controlului acesteia.

Kriging-ul, care utilizează în procesul de estimare toate valorile măsurate, este afectat în mod deosebit de ignorarea redundanței și relevanței valorilor selecției. Lipsa acestei analize afectează mai puțin metodele care utilizează un număr limitat de puncte și modele mai simple de estimare cum ar fi metoda triangulației sau metoda poligonală motiv pentru care acestea sunt preferate pentru estimări preliminare.

Un exemplu de kriging ordinar

Aplicarea kriging-ului pentru estimarea grosimii stratului impermeabil din culcușul unui acvifer freatic într-un punct p_o (**Fig.2.15**) pe baza a șase puncte în care se cunosc grosimile acestuia va clarifica toate aspectele operaționale prezentate până la această etapă.

Valorile necesare realizării estimării sunt coordonatele celor șapte puncte și valorile grosimilor măsurate în cele șase puncte (**tabelul 2.1**).

Realizarea estimării prin kriging presupune adoptarea unui model de continuitate spațială a variabilei estimate care în cazul exemplului va fi variograma sferică:

$$\tilde{\gamma}(h) = 10 \left(\frac{3}{2} \cdot \frac{h}{6} - \frac{1}{2} \cdot \frac{h^3}{6^3} \right), \quad h \in [0,6) \quad (2.61)$$

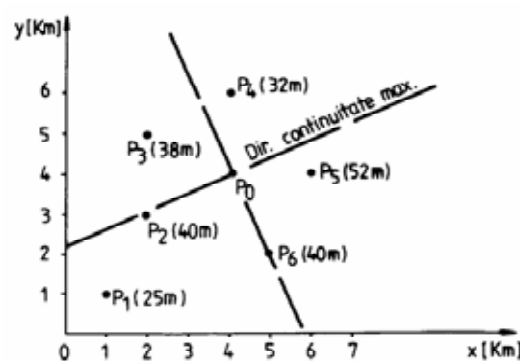


Fig.2.15. Grosimea unui acvifer sub presiune în șase piezometre.

Tabelul 2.1. Datele primare pentru kriging

Nr. crt.	x [Km]	Y [Km]	z [Km]
1	4,00	6,00	32,00
2	5,00	2,00	40,00
3	2,00	3,00	40,00
4	2,00	5,00	38,00
5	6,00	2,00	52,00
6	1,00	1,00	25,00
P	4,00	4,00	?

cu o rază de influență $r = 6$ km, valabilă în toate direcțiile. Aplicația presupune deci o structură spațială izotropă.

Utilizând modelul din ecuația (2.61) și distanțele dintre cele 7 puncte, se construiesc matricile componente ale sistemului de forma (2.49) valabil în cazul *kriging-ului punctual ordinar*:

- matricea C :

$$C = \begin{bmatrix} 10,000 & 0,434 & 1,058 & 3,739 & 0,161 & 0,000 & 1,000 \\ 0,434 & 10,000 & 1,778 & 0,327 & 7,040 & 0,434 & 1,000 \\ 1,058 & 1,778 & 10,000 & 4,320 & 0,434 & 3,739 & 1,000 \\ 3,739 & 0,327 & 4,320 & 10,000 & 0,000 & 0,434 & 1,000 \\ 0,161 & 7,040 & 0,434 & 0,000 & 10,000 & 0,000 & 1,000 \\ 0,000 & 0,434 & 0,000 & 0,434 & 0,000 & 10,000 & 1,000 \\ 1,000 & 1,000 & 1,000 & 1,000 & 1,000 & 1,000 & 0,000 \end{bmatrix} \quad (2.62)$$

- matricea D :

$$D = \begin{bmatrix} \tilde{\gamma}_{1_{p_0}} \\ \tilde{\gamma}_{2_{p_0}} \\ \tilde{\gamma}_{3_{p_0}} \\ \tilde{\gamma}_{4_{p_0}} \\ \tilde{\gamma}_{5_{p_0}} \\ \tilde{\gamma}_{6_{p_0}} \\ 1 \end{bmatrix} = \begin{bmatrix} 4,320 \\ 3,739 \\ 3,739 \\ 3,739 \\ 2,420 \\ 0,327 \\ 1,000 \end{bmatrix} \quad (2.63)$$

Prin inversarea matricii C și multiplicarea cu matricea D se obțin *ponderile* (w_i) acordate fiecărui punct de observație și *parametrul lui Lagrange* (μ):

$$w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ w_6 \\ \mu \end{bmatrix} = C^{-1} \cdot D = \begin{bmatrix} 0,352 \\ 0,309 \\ 0,262 \\ 0,133 \\ 0,018 \\ -0,074 \\ -0,108 \end{bmatrix} \quad (2.64)$$

Valoarea estimată în punctul p_o se calculează cu relația:

$$v_{p_o}^* = \sum_{i=1}^6 w_i \cdot v_i = 0,352 \cdot 32 + 0,309 \cdot 40,00 + 0,262 \cdot 40,00 + 0,133 \cdot 38,00 + \\ + 0,018 \cdot 52,0 - 0,074 \cdot 25,00 = 38,20 \text{ m} \quad (2.65)$$

iar valoarea minimă a varianței erorii de estimare cu:

$$\tilde{\sigma}_R^2 = \sum_{i=1}^n w_i \cdot \gamma_{i_{p_o}} + \mu = 0,352 \cdot 4,320 + 0,309 \cdot 3,739 + 0,262 \cdot 3,739 + \\ + 0,133 \cdot 3,739 + 0,018 \cdot 2,420 - 0,074 \cdot 0,327 - 0,108 = 5,94 \text{ m}^2 \quad (2.66)$$

Efectul parametrilor modelului de variogramă asupra estimării. Utilizarea în mod tradițional a variogramei pentru cuantificarea continuității variabilelor este argumentul pentru care în acest capitol vor fi făcute referiri la efectul parametrilor modelului de variogramă asupra rezultatelor estimărilor prin kriging.

Analiza ce va fi realizată (pe câteva exemple concrete) este o argumentare a faptului că alegerea unui model de variogramă este dificilă chiar în cazul unei rețele de probare regulate și că de ea depinde în mod esențial precizia estimărilor.

Variograma nu dă informații asupra continuității decât pentru distanțe mai mari decât distanța minimă dintre punctele de probare iar determinarea corectă a efectului de pepită și comportamentului din vecinătatea originii necesită măsurători multiple pe aceleași locații. Deoarece de cele mai multe ori nu se dispune de măsurători multiple efectul de pepită și tipul de model în vecinătatea originii, cu efect maxim asupra estimărilor, depind în mare parte de experiența practicianului.

Efectul de pepită. Valoarea mare a efectului de pepită afectează în mod negativ atât valoarea estimată prin kriging, cât și varianța erorii de estimare:

- valoarea estimată tinde spre media aritmetică a valorilor măsurate aflate în vecinătatea punctului de estimare prin egalizarea ponderilor w_i ;
- varianța erorii de estimare și, implicit, intervalul de încredere pentru valoarea estimată cresc.

În **tabelul 2.2** sunt prezentate rezultatele estimării prin kriging ordinar în punctul p_o (**Fig.2.15**) utilizându-se modele de variogramă sferică cu raza de influență $r = 6 \text{ km}$ care diferă doar prin efectul de pepită (c_o). În ecuația (2.67) sunt particularizate modelele cu efectele de pepită $c_o = 0$ și $c_o = 5$.

$$\left\{ \begin{array}{l} \gamma_1(h) = \begin{cases} 10 \left(\frac{3}{2} \cdot \frac{h}{6} - \frac{1}{2} \cdot \frac{h^3}{6^3} \right) & \text{pentru } h \in [0,6) \\ 10 & \text{pentru } h > 6 \\ 0 & \text{pentru } h = 0 \end{cases} \\ \gamma_2(h) = \begin{cases} 5 + 5 \left(\frac{3}{2} \cdot \frac{h}{6} - \frac{1}{2} \cdot \frac{h^3}{6^3} \right) & \text{pentru } h \in (0,6) \\ 10 & \text{pentru } h \geq 6 \end{cases} \end{array} \right. \quad (2.67)$$

Din analiza valorilor sintetizate în **tabelul 2.2** rezultă că o dată cu creșterea efectului de pepită de la $c_o = 0$ la $c_o = 60$:

- *amplitudinea ponderilor* (w_i) se reduce de la $A_0 = 0,378$ (corespunzătoare efectului de pepită $c_0 = 0$) la $A_{60} = 0,054$ (corespunzătoare efectului de pepită $c_0 = 60$), adică are loc o egalare a ponderilor acordate valorilor măsurate, ceea ce conduce la o valoare estimată $v^*(p_0)$ egală cu media aritmetică a valorilor măsurate;
- *varianța erorilor de estimare* crește de la $4,80 \text{ m}^2$ la $68,14 \text{ m}^2$, care reprezintă dispersia celor șase valori.

Tabelul 2.2. Rolul efectului de pepită în kriging

Nr.	Coordonate			Distanță (d_i)	Efect de pepită (Co)				
	X	Y	v_i		0	5	10	20	60
Crt.	[km]	[km]	[m]		(w _i)				
1	1	1	25	4,25	- 0,027	0,019	0,058	0,113	0,136
2	2	3	40	2,24	0,260	0,194	0,180	0,169	0,167
3	2	5	38	2,24	0,130	0,175	0,180	0,177	0,173
4	4	6	32	2,00	0,348	0,277	0,248	0,208	0,190
5	6	2	52	2,83	0,038	0,124	0,141	0,157	0,162
6	5	2	40	2,24	0,301	0,211	0,193	0,176	0,172
P ₀	4	4	?	0,00	Rezultatele estimării				
				$v^*(p_0)[m]$	38,6	38,6	38,5	38,2	38,6
				$\tilde{\sigma}_R(p_0)$	2,19	3,31	4,20	6,36	8,69
				$\tilde{\sigma}_R^2(p_0)$	4,80	11,00	17,00	40,50	68,14

Efectul de scară. Sub numele de *efect de scară* se înțelege efectul modificării valorii de plafon a variogramei asupra rezultatelor estimărilor prin kriging. Pentru același amplasament al valorilor (**Fig.2.15**) este evaluată prin kriging grosimea acviferului în p_0 utilizând două variograme de tip sferic care diferă doar prin valoarea palierului:

$$\begin{cases} \gamma_1(h) = 10 \left(\frac{3}{2} \cdot \frac{h}{6} - \frac{1}{2} \cdot \frac{h^3}{6^3} \right) \\ \gamma_2(h) = 20 \left(\frac{2}{3} \cdot \frac{h}{6} - \frac{1}{2} \cdot \frac{h^3}{6^3} \right) \end{cases}, h \in [0,6] \quad (2.68)$$

Tabelul 2.3. Efectul de scară în kriging

Nr.	Coordonate			Distanță (d_i)	Palier (C(o))	
	x	y	v_i		0	10
crt.	[km]	[km]	[m]		Ponderile (w _i)	
1	1	1	25	4,25	- 0,027	- 0,027
2	2	3	40	2,24	0,260	2,560
3	2	5	38	2,24	0,130	0,130
4	4	6	32	2,00	0,348	0,348
5	6	2	52	2,83	0,038	0,038
6	5	2	40	2,24	0,301	0,301
P ₀	4	4	?	0,00	Rezultatele	
				$V^*(p_0);m$	38,6	38,6
				$\tilde{\sigma}_R(p_0)$	2,19	3,10
				$\tilde{\sigma}_R^2(p_0)$	4,80	9,60

Rezultatele estimării sunt sintetizate în **tabelul 2.3** din analiza căruia rezultă că:

- *valorile estimate* cu cele două modele sunt identice și egale cu 38,6 m;
- *varianța erorii de estimare* este proporțională cu valoarea plafonului, raportul varianței erorilor de estimare este egal cu raportul valorilor plafonului celor două modele ($9,60/4,80 = 20/10$).

Din punct de vedere operațional modelele de variogramă cu *valori mari ale palierului* conduc la sisteme de kriging *fără soluție numerică*. Pentru realizarea estimării se utilizează un model de variogramă de același tip, dar cu un *palier redus*, operându-se după rezolvarea sistemului de kriging doar o amplificare a varianței erorii de estimare (σ_R^2) cu factorul de reducere a plafonului, valoarea estimată ($v^*(p_0)$) nefiind afectată de modificarea plafonului.

Efectul de formă. Forma variogramei, adică modelul analitic ales, influențează atât valoarea estimată cât și varianța erorii de estimare. Pentru exemplificare în tabelul 5.4 sunt prezentate rezultatele estimării prin kriging obținute cu două modele de variogramă care au același plafon ($c(0) = 10$) aceeași rază ($r = 6$ km), unul fiind de tip *gaussian* (2.69) și celalalt de tip *sferic* (2.70).

$$\gamma(h) = 10 \left[1 - \text{EXP} \left(\frac{h}{6} \right)^2 \right], h \in [0,6] \quad (2.69)$$

$$\gamma(h) = 10 \left(\frac{3}{2} \cdot \frac{h}{6} - \frac{1}{2} \cdot \frac{h^3}{6^3} \right), h \in [0,6] \quad (2.70)$$

Tabelul 2.4. Efectul de formă în kriging.

Nr. crt.	Coordonate			Distanță (d _i)	Model de variogramă	
	x [km]	y [km]	v _i [m]		Sferic	Gaussian
1	1	1	25	4,24	0,077	0,260
2	2	3	40	2,24	0,260	0,505
3	2	5	38	2,24	0,130	0,100
4	4	6	32	2,00	0,348	0,499
5	6	2	52	2,83	0,038	0,229
6	5	2	40	2,24	0,301	0,586
p ₀	4	4	?	0,00	Rezultatele estimării	
				z*(p ₀) [m]	38,6	37,40
				σ _R (p ₀)	2,19	1,17
				σ _R ² (p ₀)	4,80	1,36

Cu modelul gaussian, care indică o foarte bună continuitate a grosimii acviferului, ponderile acordate valorilor din p_2 , p_4 și p_6 , cele mai apropiate de punctul p_0 sunt cele mai mari.

Valorile p_1 , p_3 și p_5 au ponderi foarte mici, negative chiar, ele fiind *ecranate* datorită modelului ales. Această ecranare este evidentă pentru valoarea măsurată în punctul p_3 , care este ecranată de p_2 și p_4 și cea din punctul p_5 care este ecranată de valoarea din punctul p_6 .

Gradul de ecranare depinde de gradul de continuitate al modelului ales. Utilizarea modelului sferic atenuează ecranările și din cele trei valori ecranate numai cea din p_1 , care este cea mai depărtată ($d_1 = 4,24$ km), mai are pondere negativă.

Chiar cu modele lineare de variogramă care indică o continuitate mai slabă în vecinătatea originii decât cele sferice este posibil să se obțină ecranări ale unor valori (adică valori negative ale ponderilor pentru valorile respective).

Avantajul metodelor care utilizează ponderi pozitive și negative, cu sumă unitară, este că prin estimare pot conduce la valori mai mari sau mai mici decât cea mai mare, respectiv cea mai mică valoare măsurată. Toate metodele de interpolare care utilizează ponderi exclusiv pozitive conduc la valori cuprinse între valoarea minimă și maximă a eșantionului de date disponibil.

Dezavantajul metodelor care operează cu ponderi negative și pozitive este că pot conduce la valori estimate negative atunci când o valoare mare este asociată cu o pondere negativă mare. În general se operează cu parametri pozitivi: cota nivelului piezometric, grosimea acviferului, porozitatea, transmisivitatea, concentrația poluantului, debitul unei surse, etc. Pentru astfel de variabile, când estimarea prin kriging conduce la valori negative, este perfect justificat ca valoarea negativă să fie înlocuită cu valoarea zero. Pentru evitarea efectelor de ecranare, chiar în seturile de date cu o foarte bună continuitate se preferă excluderea modelelor gaussiene și parabolice chiar dacă variograma experimentală le recomandă.

Efectul de rază. Modificarea razei modelului de variogramă are o influență relativ mică asupra ponderilor acordate valorilor măsurate. Chiar dacă este redusă, această influență se resimte atât asupra valorii estimate cât și asupra varianței erorii de estimare.

Creșterea razei de influență are ca efect *apropierea statistică* a punctelor de observație. Dacă raza de influență devine mai mică decât distanța minimă dintre punctul de estimare și punctele de observație toate valorile vor primi ponderi egale în sistemul de kriging. Valoarea estimată în acest caz va fi egală cu media aritmetică a valorilor din vecinătatea punctului p_o .

În **tabelul 2.5** sunt prezentate rezultatele estimării grosimii acviferului în p_o pentru trei modele sferice de variogramă care nu diferă decât prin raza de influență. Este evidentă:

- *diferențierea ponderilor* o dată cu creșterea razei de influență. Pentru o rază de influență $r = 0,1$ km, cu mult sub distanța minimă dintre punctul de estimare (p_o) și punctele de observație, ponderile sunt egale (0,167), valoarea estimată fiind egală cu media aritmetică a celor șase valori măsurate (37,83 m);

- *reducerea varianței erorii de estimare* simultan cu creșterea razei de influență. De reținut că pentru distanțe mai mari ca raza de influență a variogramei corelația spațială dintre valori nu mai contribuie la reducerea varianței erorii de estimare.

Utilizarea în calcule a valorilor plasate la distanțe mai mari decât raza de influență a variogramei nu ameliorează rezultatul estimărilor dar nici nu-i afectează precizia. Pentru reducerea efortului de calcul și dimensionarea lui corespunzător unei precizii maxime este necesară utilizarea strategiei de selectare a punctelor prezentată în paragraful precedent.

Tabelul 2.5. Efectul de rază în kriging.

Nr crt	Coordonate			Distanță (d_i)	Raza [km]		
	x	y	v_i		0,1	6	15
	[km]	[km]	[m]		Ponderile (w_i)		
1	1	1	25	4,24	0,167	0,077	0,057
2	2	3	40	2,24	0,167	0,260	0,237
3	2	5	38	2,24	0,167	0,130	0,134
4	4	6	32	2,00	0,167	0,348	0,366
5	6	2	52	2,83	0,167	0,038	0,069
6	5	2	40	2,24	0,167	0,301	0,281
P_0	4	4	?	0,00	Rezultatele estimării		
				$z^*(p_0)[m]$	37,83	38,60	38,70
				$\sigma_R(p_0)$	3,42	2,19	1,35
				$\sigma_R^2(p_0)$	11,70	4,80	1,81

Efectul de anizotropie. În toate exemplele prezentate au fost utilizate modele de variogramă izotrope care ignoră rolul direcției în studiul continuității. Harta conturală a unor astfel de modele este formată din cercuri concentrice spre deosebire de modelele anizotrope.

Utilizând un model de variogramă anizotrop de tip geometric (același model și palier în toate direcțiile și raze diferite, **Fig.1.55a**) cu direcțiile de anizotropie N60°E (continuitate maximă) și N30°V (continuitate minimă) razele de influență $r_1 = 15$ km, respectiv $r_2 = 5$ km estimările din p_0 conduc la rezultatele din **tabelul 2.6**.

Tabelul 2.6. Efectul de anizotropie în kriging.

Nr. crt.	Coordonate			Distanță (d_i)	N 60°E	N 30°V	N 60°E	N 30°V
	x	y	v_i		Raza de influență [km]			
	[km]	[km]	[m]		15	5	15	5
					Ponderile (w_i)			
1	1	1	25	4,24	0,015		- 0,057	
2	2	3	40	2,24	0,638		0,237	
3	2	5	38	2,24	- 0,046		0,134	
4	4	6	32	2,00	0,236		0,366	
5	6	2	52	2,83	- 0,008		0,069	
6	5	2	40	2,24	0,165		0,281	
P_0	4	4	?	0,00	Rezultatele estimării			
				$z^*(p_0)[m]$	37,9		38,7	
				$\sigma_R(p_0)$	1,87		1,35	
				$\sigma_R^2(p_0)$	3,50		1,81	

În **tabelul 2.6** sunt prezentate rezultatele obținute cu *modelul sferic izotrop* cu o rază $r = 15$ km și cu *modelul sferic anizotrop* care evidențiază creșterea considerabilă a ponderilor valorilor plasate pe direcția de continuitate maximă și reducerea celor plasate pe direcția de continuitate redusă.

Valorile din punctele p_3 și p_5 sunt puternic ecranate fiind plasate pe direcția de continuitate minimă și având în cazul modelului anizotrop ponderi negative, în timp

ce ponderea valorii din punctul p_2 , plasat pe direcția de continuitate maximă, are o creștere spectaculoasă de la 0,237 în cazul modelului izotrop la 0,638 în cazul modelului anizotrop utilizat.

COMENTARIU

Analiza influenței parametrilor variogramei asupra valorilor estimate și a erorilor de estimare argumentează importanța etapei de analiză a variabilității caracteristicilor studiate pentru estimarea distribuției spațiale.

Trebuie reținut că asupra valorii estimate efectul important este dat de raza de influență și efectul de pepită al variogramei:

- *cu cât raza de influență a variogramei este mai mare cu atât valoarea estimată este mai aproape de cea reală;*
- *cu cât efectul de pepită al modelului de variogramă este mai mare cu atât valoarea estimată într-un punct oarecare se apropie de media aritmetică a valorilor măsurate, îndepărtându-se sau... apropiindu-se... (nimeni nu mai poate ști!), de valoarea reală din acel punct.*

Valoarea palierului variogramei este direct proporțională cu mărimea erorii care afectează operațiunea de estimare. Valoarea estimată nefiind afectată de mărimea valorii palierului variogramei, un anumit model de variogramă poate fi utilizat pentru proiectarea rețelei optime de explorare a unei structuri spațiale. Această proprietate este utilizată în metoda punctului fictiv pe baza căreia se ameliorează performanțele rețelelor de cercetare.

Atenție la relevanța datelor pentru obiectivul estimării!

Corectitudinea estimării prin kriging este condiționată nu numai de corectitudinea aplicării metodologiei ci și de clarviziunea celui care-și selecționează datele dincolo de criteriile statistice.

Este determinantă pentru realismul estimării selectarea datelor care reflectă distribuția spațială rezultată din desfășurarea procesului care constituie obiectul cercetării. Metodologia geostatistică operează asupra valorilor pe care noi le selectăm și relevă caracteristici ale acestor valori. Interpretarea ne aparține și este cu atât mai simplă cu cât complexitatea procesului studiat este mai redusă.

Rețineți!

Chiar rezultatele estimării distribuției spațiale pot să ne semnaleze necesitatea revenirii la etapa de analiză a variabilității pentru a separa efectele unor procese care se suprapun în timp și spațiu pe zona cercetată.

Identificarea unor zone cu variabilitate extremă de exemplu, poate indica necesitatea separării selecției de date prelucrate pe baza geometriei zonelor respective și revenirea la prima etapă de analiză a variabilității.

A21. Realizarea secțiunilor litologice

Să se construiască secțiunea litologică de-a lungul unui aliniament rectiliniu de 40 m pe care sunt amplasate cinci foraje de explorare de 35-40 m adâncime. Succesiunea litologică traversată este formată din argile și nisipuri (Fig.2.16).

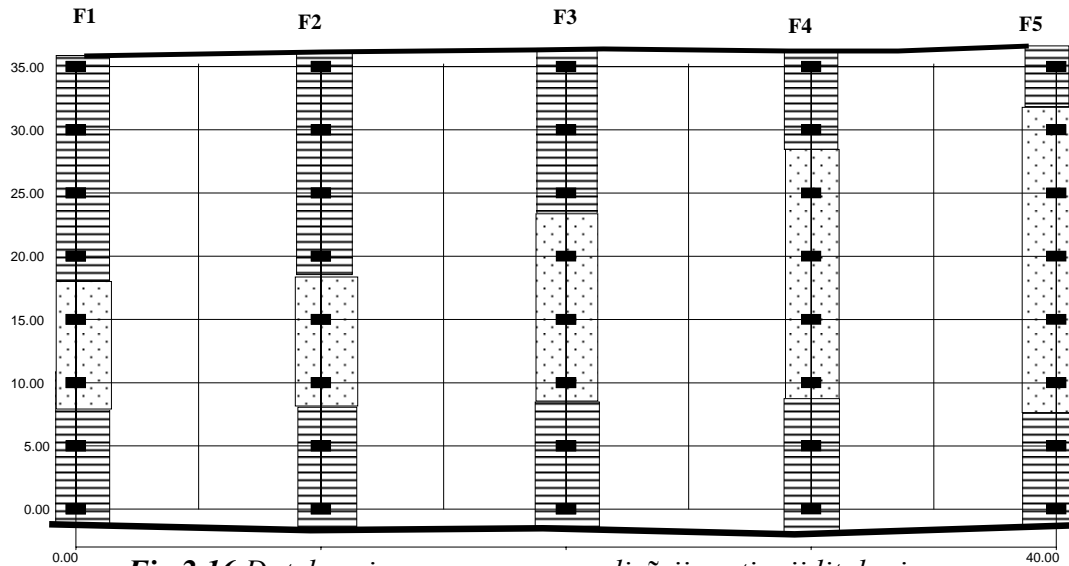


Fig.2.16. Datele primare necesare realizării secțiunii litologice

Rezolvare:

O secțiune geologică presupune integrarea unui mare volum de informații de natură calitativă și cantitativă: litologia formațiunilor, tipul bazinului de sedimentare în care au fost depuse formațiunile, istoria tectonică a zonei etc.

Datele necesare realizării secțiunilor litologice, în care se reflectă acțiunea tuturor factorilor care au contribuit la formarea structurii geologice, se obțin din aflorimente și prin lucrări de explorare geologică (foraje și lucrări miniere de suprafață și subterane). Aceste date constau în succesiuni litologice și orientări spațiale ale limitelor de separație a formațiunilor.

Tehnic vorbind, construirea secțiunii se rezumă la “corelarea” formațiunilor identificate, corelare realizată într-un anumit *stil structural* determinat de “*ambianța geologică*”. Corelarea este un fel de interpolare a formațiunilor identificate care implică legități determinate de tipul bazinului de sedimentare, evoluția tectonică etc.

Metodele geostatistice care sunt utilizate pentru realizarea secțiunilor litologice sunt cele de *estimare punctuală* și anume *kriging-ul punctual ordinar* în varianta lui “*indicator*”. *Kriging-ul indicator* poartă această denumire pentru că operează, la fel ca și *variograma indicator*, asupra a două valori: *unu* și *zero*. Tehnica de prelucrare este identică cu a *kriging-ului ordinar punctual*.

Kriging-ul indicator permite integrarea informației primare cu elemente suplimentare de tipul: *orientarea direcțiilor de anizotropie*, *raportul de anizotropie*, *gradul de certitudine al determinărilor* (exprimat prin abaterea standard a valorilor) etc.

Realizarea secțiunii litologice presupune șapte etape de prelucrare din care primele patru, preliminară, au fost parcurse în aplicația A14 și au condus la

determinarea caracteristicilor principale ale structurii și la identificarea legităților de variație spațială a argilei și nisipului:

- analiza distribuției spațiale, atât pe *orizontală* cât și pe *verticală*, a celor doi litotipi prezenți în cele cinci foraje (argilă și nisip) a identificat trei unități structurale: două pentru argilă și una pentru nisip.
- *unitatea structurală superioară a argilei* are o *variogramă indicatoare izotropă* de tip *sferic* cu parametrii:
 - *efectul de pepită* = 0,045;
 - *palierul* = 0,25;
 - *raza de influență* = 23 m.
- *unitatea structurală a nisipului* se manifestă de asemenea *izotrop*, variograma indicatoare fiind tot de tip *sferic* cu parametrii:
 - *efectul de pepită* = 0,08;
 - *palierul* = 0,2;
 - *raza de influență* = 20 m.

Pentru parcurgerea următoarelor două etape *datele* necesare estimării distribuției spațiale a argilei și nisipului sunt:

- valorile din **tabelul 1.22** (*coordonatele punctelor de observație și codurile numerice ale litotipurilor* organizate în fișier de tip Geo-EAS sau tip SURFER);
- parametrii modelului de variogramă pentru fiecare litotip (*tipul modelului, efectul de pepită, palierul, raza*).

Realizarea secțiunii litologice cu cei doi litotipi presupune următoarele prelucrări, numerotarea lor fiind în continuarea celor de la aplicația *A14* (pag.91):

5. Alegerea probabilității minime de identificare a litotipurilor în secțiune.

În secțiunea pe care intenționăm să o realizăm, probabilitatea de apariție a argilei poate lua două valori extreme:

- *zero* în punctele în care a fost identificat nisipul;
- *unu* în punctele în care a fost identificată argila.

Pentru celelalte puncte, valorile probabilității de apariție a argilei vor fi cuprinse între aceste două valori:

- *valori mai apropiate de unu* în vecinătatea punctelor în care a fost identificată *argila*;
- *valori mai apropiate de zero* în vecinătatea punctelor în care a fost identificat *nisipul*.

Variația probabilității de apariție a argilei în funcție de distanță este cuantificată în modelul variogramei indicatoare definitivat în etapa a patra.

Realizarea secțiunii presupune alegerea unei *valori minime a probabilității de apariție a argilei*. În punctele în care valoarea probabilității de apariție a argilei depășește această valoare minimă aleasă se consideră că este prezentă argila.

Alegerea valorii minime a probabilității de apariție a argilei se face în funcție de gradul de certitudine pe care vrem să-l avem pentru secțiunea realizată. Cu cât valoarea minimă aleasă pentru această probabilitate este mai mare cu atât domeniile spațiale de apariție a argilei se restrâng. De obicei se alege un nivel de probabilitate $> 0,5$ pentru a vizualiza în secțiune domeniile spațiale pe care probabilitatea de apariție a argilei studiate este mai mare decât cea de absență a acesteia.

6.Construirea secțiunilor cu izolinii pentru distribuția *probabilității de apariție a argilei (Fig.2.17a)* și nisipului (*Fig.2.17b*). Cele două secțiuni se construiesc separat

pentru argilă și nisip în scopul verificării reprezentativității valorii minime a probabilității utilizate pentru definitivarea secțiunii litologice.

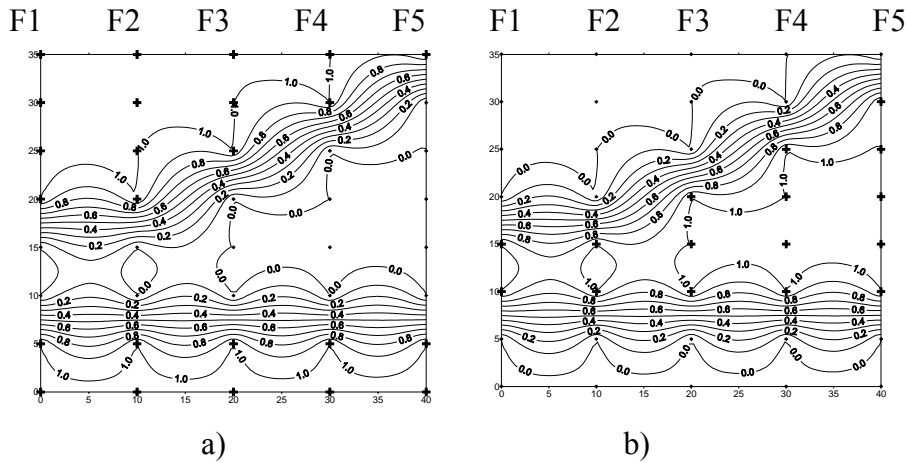


Fig.2.17. Secțiunile cu izoprobabilități de apariție ale argilei (a) și nisipului (b).

Cu cât valoarea minimă de probabilitate aleasă este mai mare cu atât domeniul spațial de extindere a litotipului studiat este mai restrâns. Pentru exemplificare sunt conturate domeniile de extindere spațială a argilei pentru trei valori ale probabilității minime de apariție a argilei: $p_1 = 0,3$ (**Fig.2.18a**); $p_2 = 0,7$ (**Fig.2.18b**) și $p_3 = 0,9$ (**Fig.2.18c**).

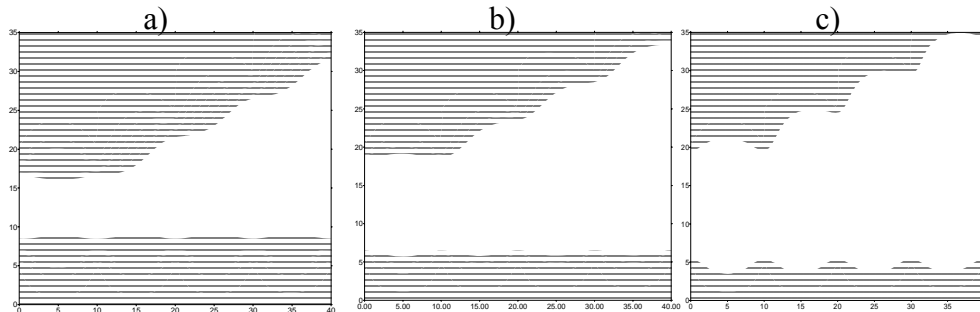


Fig.2.18. Secțiuni cu izoprobabilități de apariție a argilei de $p_1 > 30\%$ (a), $p_2 > 70\%$ (b), $p_3 > 90\%$ (c) (domeniul de extindere al argilei este hașurat cu linii orizontale).

Domeniul nehașurat corespunde unor probabilități de apariție a argilei mai mici decât probabilitățile pentru care sunt construite cele trei secțiuni. În cazul studiat, zonele nehașurate corespund unor probabilități de apariție a nisipului mai mari decât cele de apariție a argilei. Astfel, domeniul nehașurat din **Fig.2.18a** corespunde zonei în care probabilitatea de apariție a nisipului este mai mare de 30%. Acest lucru nu mai este valabil când în secțiune sunt prezente mai mult de două litologii.

Construirea secțiunilor cu izolinii de probabilități începe cu litotipul având frecvența cea mai mare din *histograma nominală* (**Fig.1.62**). Frecvența maximă de apariție a unui litotip dovedește gradul avansat de cunoaștere a distribuției acestuia în comparație cu ceilalți. Acest lucru se reflectă și în variograma indicatoare a cărei formă este mai precis conturată pentru litotipii cu frecvență mai mare de apariție.

7. Suprapunerea secțiunilor corespunzătoare extinderii nisipului și argilei pentru nivelul de probabilitate ales.

Pentru nivelul comun de probabilitate minimă $p = 0,5$, secțiunea litologică obținută prin suprapunerea celor realizate separat pentru cei doi litotipi indică domeniile spațiale în care prezența fiecărui litotip este mai mare de 0,5 (**Fig.2.19**).

Numărul litotipilor nu modifică metodologia de lucru, analiza și estimarea fiecăruia desfășurându-se separat. În final se suprapun secțiunile realizate la un nivel comun de probabilitate pentru toți litotipii.

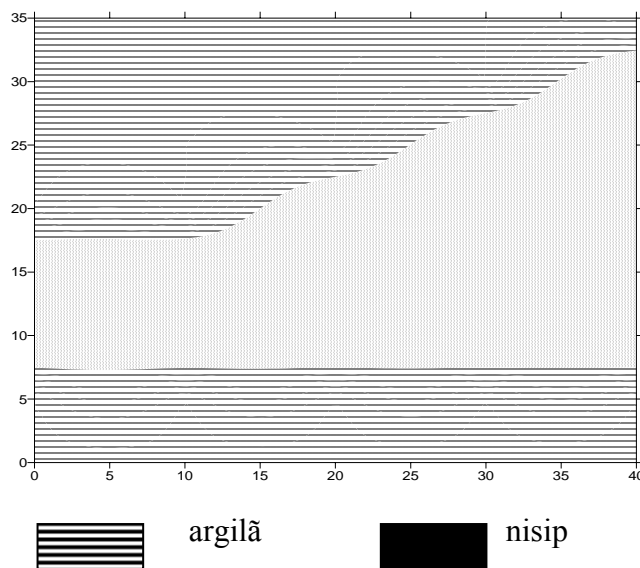


Fig.2.19. Secțiunea litologică pentru un nivel de probabilitate $p > 50\%$.

Dacă nivelul comun de probabilitate minimă este mai mare de 0,5 atunci la limita de separație dintre doi litotipi rămân suprafețe neacoperite în care erorile de estimare sunt mai mari decât în

restul hărții. Pentru un nivel comun de probabilitate minimă $p = 0,8$ se dezvoltă o bandă la limita dintre cei doi litotipi în care probabilitatea de apariție a argilei și a nisipului este mai mică de 0,8 (**Fig.2.20**).

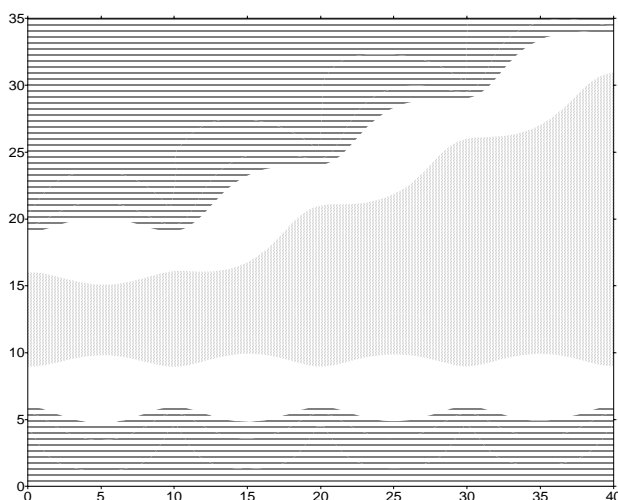


Fig.2.20. Secțiune litologică pentru $p > 80\%$ pentru ambii litotipi.

Gradul de certitudine al acestor secțiuni litologice se estimează prin intermediul secțiunilor cu izolinii al varianței erorii de estimare sau al intervalelor de încredere pentru probabilitățile estimate.

COMENTARIU

Pentru exersarea metodologiei vă ofer două seturi de date și câte unul din cele mai probabile rezultate care se pot obține. Nu trebuie să fiți îngrijorați dacă nu veți obține exact același rezultat cu cel prezentat în figurile anexate tabelor de date.

Pe baza acelorași date, zece geologi fac de cele mai multe ori zece hărți diferite!

Diferențele dintre secțiuni sunt determinate fie de alegerea modelului de variogramă fie de probabilitatea aleasă ca valoare minimă pentru reprezentarea diferiților litotipi din secțiuni.

Tabelul 2.7. Date codificate pentru exercițiul 1 (C - calcar; A - argilă; G - gresie)

nr	X	y	C	A	G	nr	x	y	C	A	G
1	0	0	1	0	0	41	4	6	1	0	0
2	0	1	1	0	0	42	4	7	1	0	0
3	0	2	1	0	0	43	4	8	0	1	0
4	0	3	1	0	0	44	4	9	0	1	0
5	0	4	1	0	0	45	4	10	0	1	0
6	0	5	1	0	0	46	4	11	0	1	0
7	0	6	0	1	0	47	4	12	0	1	0
8	0	7	0	1	0	48	4	13	0	1	0
9	0	8	0	1	0	49	4	14	0	1	0
10	0	9	0	1	0	50	4	15	0	1	0
11	0	10	0	1	0	51	4	16	0	1	0
12	0	11	0	1	0	52	4	8	0	1	0
13	0	12	0	1	0	53	4	9	0	1	0
14	0	13	0	1	0	54	4	10	0	1	0
15	0	14	0	1	0	55	4	11	0	1	0
16	0	15	0	1	0	56	4	12	0	1	0
17	0	16	0	1	0	57	4	13	0	1	0
18	2	0	1	0	0	58	6	0	1	0	0
19	2	1	1	0	0	59	6	1	1	0	0
20	2	2	1	0	0	60	6	2	1	0	0
21	2	3	1	0	0	61	6	3	1	0	0
22	2	4	1	0	0	62	6	4	1	0	0
23	2	5	1	0	0	63	6	5	1	0	0
24	2	6	1	0	0	64	6	6	1	0	0
25	2	7	0	1	0	65	6	7	1	0	0
26	2	8	0	1	0	66	6	8	1	0	0
27	2	9	0	1	0	67	6	9	0	1	0
28	2	10	0	1	0	68	6	10	0	1	0
29	2	11	0	1	0	69	6	11	0	1	0
30	2	12	0	1	0	70	6	12	0	1	0
31	2	13	0	1	0	71	6	13	0	1	0
32	2	14	0	1	0	72	6	14	0	1	0
33	2	15	0	1	0	73	6	15	0	1	0
34	2	16	0	1	0	74	6	16	0	1	0
35	4	0	1	0	0	75	8	0	1	0	0
36	4	1	1	0	0	76	8	1	1	0	0
37	4	2	1	0	0	77	8	2	1	0	0
38	4	3	1	0	0	78	8	3	1	0	0
39	4	4	1	0	0	79	8	4	1	0	0
40	4	5	1	0	0	80	8	5	1	0	0

Exercițiul 1

Să se realizeze secțiunea litologică pe baza informațiilor obținute dintr-un număr de 14 foraje probate din metru în metru.

Forajele au traversat o secvență constituită din trei litotipi (A - argilă, G - gresie; C - calcar) iar datele codificate ($n = 229$ de puncte) sunt grupate în **tabelul 2.7**.

Soluție posibilă

Secțiunea litologică pentru o probabilitate comună minimă $p=0,5$ este prezentată în **Fig.2.21**.

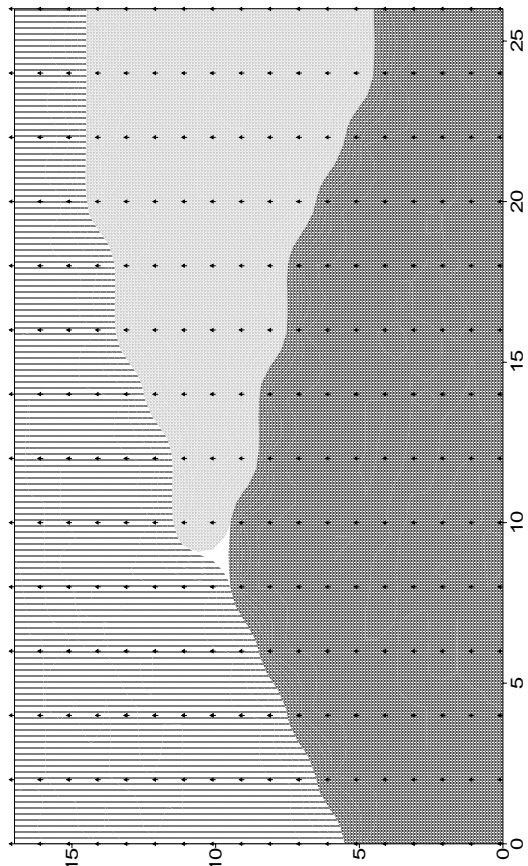


Fig.2.21. Secțiune litologică cu trei litotipi.

Tabelul 2.7(continuare)

nr	x	y	C	A	G	nr	x	y	C	A	G	nr	x	y	C	A	G	nr	x	y	C	A	G	nr	x	y	C	A	G
81	8	6	1	0	0	121	12	12	0	1	0	161	16	16	0	1	0	201	22	2	1	0	0	230	24	13	0	0	1
82	8	7	1	0	0	122	12	13	0	1	0	162	16	17	0	1	0	202	22	3	1	0	0	231	24	14	0	0	1
83	8	8	1	0	0	123	12	14	0	1	0	163	18	0	1	0	0	203	22	4	1	0	0	232	24	15	0	1	0
84	8	9	1	0	0	124	12	15	0	1	0	164	18	1	1	0	0	204	22	5	1	0	0	233	24	16	0	1	0
85	8	10	0	1	0	125	12	16	0	1	0	165	18	2	1	0	0	205	22	6	0	0	1	234	24	17	0	1	0
86	8	11	0	1	0	126	12	17	0	1	0	166	18	3	1	0	0	206	22	7	0	0	1	235	26	0	1	0	0
87	8	12	0	1	0	127	14	0	1	0	0	167	18	4	1	0	0	207	22	8	0	0	1	236	26	1	1	0	0
88	8	13	0	1	0	128	14	1	1	0	0	168	18	5	1	0	0	208	22	9	0	0	1	237	26	2	1	0	0
89	8	14	0	1	0	129	14	2	1	0	0	169	18	6	1	0	0	209	22	10	0	0	1	238	26	3	1	0	0
90	8	15	0	1	0	130	14	3	1	0	0	170	18	7	1	0	0	210	22	11	0	0	1	239	26	4	1	0	0
91	8	16	0	1	0	131	14	4	1	0	0	171	18	8	0	0	1	211	22	12	0	0	1	240	26	5	0	0	1
92	10	0	1	0	0	132	14	5	1	0	0	172	18	9	0	0	1	212	22	13	0	0	1	241	26	6	0	0	1
93	10	1	1	0	0	133	14	6	1	0	0	173	18	10	0	0	1	213	22	14	0	0	1	242	26	7	0	0	1
94	10	2	1	0	0	134	14	7	1	0	0	174	18	11	0	0	1	214	22	15	0	1	0	243	26	8	0	0	1
95	10	3	1	0	0	135	14	8	1	0	0	175	18	12	0	0	1	215	22	16	0	1	0	244	26	9	0	0	1
96	10	4	1	0	0	136	14	9	0	0	1	176	18	13	0	0	1	216	22	17	0	1	0	245	26	10	0	0	1
97	10	5	1	0	0	137	14	10	0	0	1	177	18	14	0	1	0	217	24	0	1	0	0	246	26	11	0	0	1
98	10	6	1	0	0	138	14	11	0	0	1	178	18	15	0	1	0	218	24	1	1	0	0	247	26	12	0	0	1
99	10	7	1	0	0	139	14	12	0	0	1	179	18	16	0	1	0	219	24	2	1	0	0	248	26	13	0	0	1
100	10	8	1	0	0	140	14	13	0	1	0	180	18	17	0	1	0	220	24	3	1	0	0	249	26	14	0	0	1
101	10	9	1	0	0	141	14	14	0	1	0	181	20	0	1	0	0	221	24	4	1	0	0	250	26	15	0	1	0
102	10	10	0	0	1	142	14	15	0	1	0	182	20	1	1	0	0	222	24	5	0	0	1	251	26	16	0	1	0
103	10	11	0	0	1	143	14	16	0	1	0	183	20	2	1	0	0	223	24	6	0	0	1	252	26	17	0	1	0
104	10	12	0	1	0	144	14	17	0	1	0	184	20	3	1	0	0	224	24	7	0	0	1	253	0	17	0	1	0
105	10	13	0	1	0	145	16	0	1	0	0	185	20	4	1	0	0	225	24	8	0	0	1	254	2	17	0	1	0
106	10	14	0	1	0	146	16	1	1	0	0	186	20	5	1	0	0	226	24	9	0	0	1	255	4	17	0	1	0
107	10	15	0	1	0	147	16	2	1	0	0	187	20	6	1	0	0	227	24	10	0	0	1	256	6	17	0	1	0
108	10	16	0	1	0	148	16	3	1	0	0	188	20	7	0	0	1	228	24	11	0	0	1	257	8	17	0	1	0
109	12	0	1	0	0	149	16	4	1	0	0	189	20	8	0	0	1	229	24	12	0	0	1	258	10	17	0	1	0
110	12	1	1	0	0	150	16	5	1	0	0	190	20	9	0	0	1												
111	12	2	1	0	0	151	16	6	1	0	0	191	20	10	0	0	1												
112	12	3	1	0	0	152	16	7	1	0	0	192	20	11	0	0	1												
113	12	4	1	0	0	153	16	8	0	0	1	193	20	12	0	0	1												
114	12	5	1	0	0	154	16	9	0	0	1	194	20	13	0	0	1												
115	12	6	1	0	0	155	16	10	0	0	1	195	20	14	0	0	1												
116	12	7	1	0	0	156	16	11	0	0	1	196	20	15	0	1	0												
117	12	8	1	0	0	157	16	12	0	0	1	197	20	16	0	1	0												
118	12	9	0	0	1	158	16	13	0	0	1	198	20	17	0	1	0												
119	12	10	0	0	1	159	16	14	0	1	0	199	22	0	1	0	0												
120	12	11	0	0	1	160	16	15	0	1	0	200	22	1	1	0	0												

Exercițiul 2

Să se realizeze secțiunea litologică de-a lungul unui front de captare pentru ape subterane pe baza datelor codificate din tabelul 2.8.

Pentru realizarea captării de apă din acviferul freatic, pentru municipiul Satu-Mare, s-au executat 28 de foraje de 120 m adâncime, plasate echidistant (e=250m), pe un aliniament rectiliniu. Forajele au traversat două tipuri litologice distincte: argilă și nisip.

Soluție probabilă

Secțiunea cea mai probabilă este cea din **Fig.2.22** realizată pentru probabilitatea comună minimă $p = 0,5$. (Atenție, structura are o puternică anizotropie cu direcția de variabilitate maximă pe verticala secțiunii și raportul razelor variogramelor pe cele două direcții $r_a = 1/3$.)

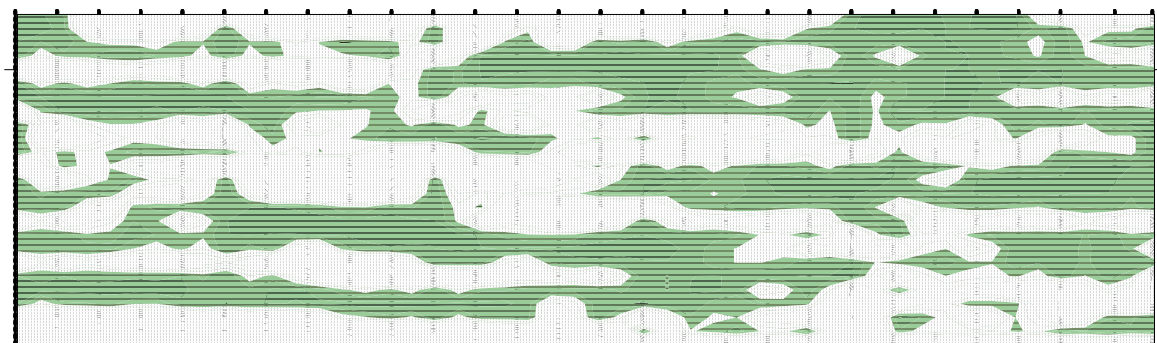


Fig.2.22. Secțiunea litologică pe traseul frontului de captare Satu Mare.

Proportional cu numărul de puncte de observație crește precizia secțiunilor realizate dar și efortul achiziționării și codificării datelor. Metodele automate de procesare a imaginilor sunt obligatoriu de utilizat pentru discretizarea și codificarea datelor necesare realizării secțiunilor și hărților litologice. În studiul **aflorimentelor fotografia** constituie un instrument important în **achiziția datelor** iar prelucrarea ei automată crește eficiența prelucrării. Calitatea materialelor utilizate și a tehnicii de prelucrare (granulația filmului și hârtiei fotografice, rezoluția scanner-ului, performanțele programelor de procesare a imaginii) influențează, în funcție de scara de prelucrare, reprezentativitatea soluțiilor obținute.

Nici unul din aceste aspecte nu trebuie neglijat în realizarea prin metode geostatistice a unei secțiuni sau hărți litologice deoarece totul se reflectă în precizia acestora!

Tehnica geostatistică de realizare a secțiunilor litologice nu poate cuantifica toate elementele care conduc la realizarea unei secțiuni corecte. Este necesar ca această tehnică să fie utilizată de un geolog care cunoaște caracteristicile fundamentale ale proceselor care pot conduce la realizarea succesiunilor litologice studiate. Metodele geostatistice operează corect doar în domenii omogene din punct de vedere al variabilității spațiale iar limitele acestor domenii trebuie, de cele mai multe ori, precizate de cel care le aplică.

Să nu vă imaginați că doar prin însușirea tehnicii geostatistice de prelucrare veți putea realiza o secțiune geologică adevărată!

Să nu desconsiderați din acest motiv modelele geostatistice deoarece sunt situații în care fără utilizarea acestora, secțiunile construite sunt doar povești frumoase în care imaginația are o pondere de... 100% .

A22. Evaluarea distribuției unei variabile staționare.

Să se evalueze distribuția durtății temporare a apelor subterane din acviferul de medie adâncime în zona captărilor orașului Slatina pe baza valorilor determinate în probele de apă prelevate din 36 de foraje de captare (tabelul 2.9).

Rezolvare:

Duritatea temporară este o importantă caracteristică chimică a apelor subterane. Ea reprezintă conținutul apei în săruri instabile de calciu și magneziu care se depun în anumite condiții sub formă de “piatră” în instalațiile unde apa este utilizată și sub formă de “calculi” în organismul celor care o beau. Acestea sunt două din motivele pentru care trebuie cunoscută distribuția spațială a durtății temporare a apei subterane care este captată pentru diferite utilizări.

Duritatea temporară se exprimă în mod uzual în grade germane ($1\text{mval/litruCa(Mg)}=1^\circ\text{h}$). Cele 36 de valori ale durtății temporare au o repartiție cvasinormală cu un coeficient de asimetrie $\beta_3 = 0,70$ (Fig.2.23).

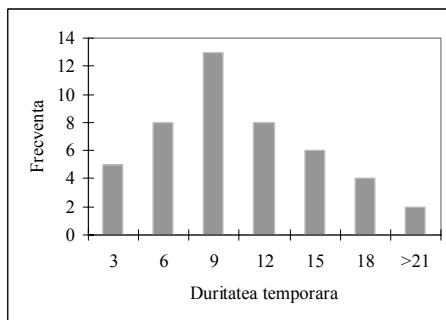


Fig.2.23 Histograma valorilor durtății temporare a apei

Erorile de supraestimare rezultate din această asimetrie nu depășesc 10% din valoarea durtăților estimate astfel încât se poate realiza estimarea distribuției spațiale cu valorile originale fără să fie absolut necesară normalizarea lor.

Analiza variografică a structurii spațiale a durtății temporare conduce la următoarele rezultate:

- distribuția durtății temporare a acviferului de medie adâncime, în zona captărilor cu apă subterană ale orașului Slatina, este anizotropă, direcția de variabilitate minimă (continuitate maximă) fiind $N15^\circ V$ (Fig.2.24);

Tabelul 2.9 Duritatea temporară

Nr. crt.	X [m]	Y [m]	D temp [°h]
1	1157.59	3991.50	6.27
2	1125.43	3911.56	7.39
3	1082.10	3181.20	5.60
4	1095.64	3127.32	10.53
5	1095.96	2607.39	12.04
6	1093.16	2551.72	7.16
7	1866.49	4033.98	15.01
8	1713.27	4002.80	13.72
9	1694.38	3948.41	6.50
10	1484.26	3544.90	9.24
11	1492.23	3140.14	7.17
12	1468.23	3095.44	12.88
13	1356.64	3093.39	8.40
14	1442.44	3069.14	12.88
15	1407.50	2737.92	8.40
16	1389.65	2542.05	9.52
17	1385.31	2500.68	6.05
18	1329.15	2296.12	7.61
19	1327.65	2081.10	7.84
20	1053.13	2036.37	15.96
21	525.00	2273.29	17.80
22	830.19	2061.37	9.41
23	851.65	2020.01	19.04
24	2120.36	1719.59	21.84
25	1817.00	1591.10	5.60
26	1907.67	1462.92	7.84
27	1935.77	1415.68	3.36
28	1161.47	1575.19	13.44
29	1147.69	1490.66	7.39
30	1228.66	1331.83	17.60
31	1239.39	1274.89	6.27
32	1821.44	806.61	15.45
33	1819.14	759.37	5.15
34	1623.79	739.93	17.36
35	1793.14	411.04	4.48
36	1797.23	351.28	4.92

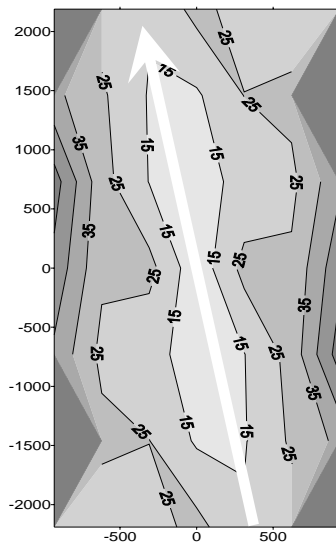


Fig.2.24 Variograma de suprafață a durtății temporare

- modelul variogramei structurii este de tip sferic (**Fig.2.25**) cu parametrii:
 - efect de pepită = 10;
 - palier = 15;
 - razele de influență :
 - $r_1 = 1200$ m pe direcția N15°V;
 - $r_2 = 600$ m pe direcția N75°E.

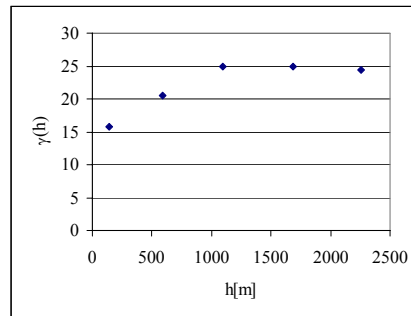


Fig.2.25 Variograma omnidirecțională a durtății temporare

Estimarea distribuției spațiale a durtății temporare s-a realizat prin *kriging punctual ordinar* (deoarece structura este staționară, modelul de variogramă fiind cu palier). Au fost utilizate 3901 puncte de calcul amplasate într-o rețea rectangulară cu 83 de rânduri și 47 de coloane.

În zona mediană a hărții cu izolinii obținute (**Fig.2.26**) sunt figurate lacurile de acumulare realizate în această zonă pe râul Olt.

Se remarcă un paralelism între direcția de curgere a Oltului și liniile de egală valoare a durtății temporare.

Această caracteristică sugerează existența unei comunicări între acviferul de medie adâncime și râul Olt, comunicare care face ca

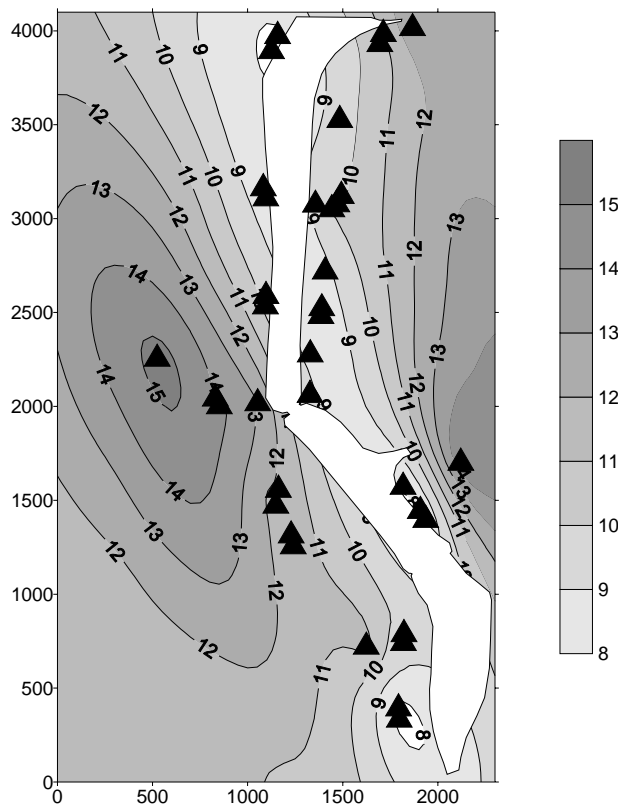


Fig.2.26 Distribuția durtății temporare a acviferului de medie adâncime în zona orașului Slatina

duritatea temporară să fie cu atât mai mare cu cât ne situăm la o distanță mai mare de Olt.

COMENTARIU

Estimarea distribuției spațiale a unei variabile este o operațiune extrem de simplă. Toate dificultățile care stau în fața estimării distribuției spațiale sunt depășite în etapele de analiză a variabilității caracteristicilor studiate.

În etapa estimării distribuției spațiale parametrii care trebuie stabiliți vizează numai două aspecte:

- *gradul de detaliere la care este făcută estimarea;*
- *claritatea și expresivitatea vizualizării grafice a rezultatelor.*

Gradul de detaliere la care este realizată estimarea distribuției spațiale este determinat de densitatea punctelor în care se face estimarea valorii caracteristicii studiate. Programele de calcul care realizează aceste estimări solicită parametrii unei rețele rectangulare cu care se acoperă suprafața pe care se face estimarea:

- *originea rețelei rectangulare ($O(x_0, y_0)$);*
- *numărul de rânduri ale rețelei (ny);*
- *numărul de coloane ale rețelei (nx);*
- *distanța dintre coloane (Δx);*
- *distanța dintre rânduri (Δy).*

Estimarea valorii caracteristicii studiate se face în fiecare celulă rectangulară a rețelei și este cu atât mai detaliată cu cât numărul de celule este mai mare.

Creșterea gradului de detaliere nu conduce la creșterea preciziei estimării !

Precizia cu care este realizată estimarea distribuției spațiale depinde în principal de numărul de puncte de observație în care se cunoaște valoarea caracteristicii ce se estimează.

Claritatea și expresivitatea vizualizării grafice a rezultatelor estimării distribuției spațiale depind de opțiunile și bunul gust al autorului hărții. Sunt câteva principii pe care este bine să le respecti când finalizezi harta cu izolinii a unei distribuții spațiale:

- *echidistanța între izolinii să permită citirea etichetelor inserate;*
- *numărul elementelor secundare suprapuse pe hartă (elemente topografice, semne convenționale etc.) să fie redus;*
- *să existe elemente care să permită verificarea corectitudinii morfologiei izoliniilor (poziția punctelor de observație și valoarea măsurată, în limita unei densități care să nu producă suprapuneri)*

Aspectul general al hărții trebuie să evidențieze scopul pentru care a fost realizată!

Harta trebuie să fie clară și... argumentată! Ea este realizată pentru a demonstra ceva iar demonstrația trebuie să fie convingătoare.

Nu încercați să realizați hărți pentru a ascunde adevărul sau insuficienta cunoaștere!

Dacă ai de ascuns ceva nu recurge niciodată la reprezentări grafice!

Se vede imediat!

2.2.2. KRIGING PUNCTUAL UNIVERSAL

Toate sistemele de kriging prezentate până aici presupun pentru variabila studiată un model de funcție aleatoare staționară sau cvasistaționară în vecinătatea punctului în care se face estimarea.

Deseori, pentru anumite variabile, se identifică o tendință zonală (speranța matematică nu este staționară) iar informațiile disponibile nu sunt suficient de dense pentru a lua în considerare vecinătăți cvasistaționare. Este cazul suprafețelor piezometrice ale acviferelor cu dinamică activă în care este prezentă o tendință regională.

Aplicarea *kriging-ului punctual ordinar (staționar)* în prezența unei tendințe va conduce în mod sistematic la supraevaluări ale variabilei studiate. Pentru eliminarea erorilor de estimare trebuie să se țină seama de prezența și forma acestei tendințe.

Kriging-ul punctual universal (sau kriging-ul nedeviat de ordinul k) furnizează un estimator nedeviat ce ține seama de prezența tendinței cu condiția cunoașterii formei acesteia și covarianței sau variogramei modelului funcției aleatoare nestaționare a variabilei.

Forma tendinței regionale. În cazul unui model de funcție aleatoare nestaționară, prin definiție, tendința variabilei regionalizate este speranța matematică nestaționară:

$$E\{V(p)\} = m(p) \quad (2.71)$$

Funcția aleatoare poate fi descompusă într-o tendință ($m(p)$) și un termen rezidual $Y(p)$ staționar sau nestaționar dar cu speranța matematică nulă:

$$V(p) = m(p) + Y(p) \text{ cu } E\{Y(p)\} = 0 \quad (2.72)$$

Tendința $m(p)$ reprezintă variația regulată a funcției aleatoare la scara distribuției punctelor de observație, iar reziduul ($Y(p)$) fluctuațiile aleatoare, dar regionalizate, de o parte și de alta a tendinței (**Fig. 2.27**).

Forma tendinței este o combinație lineară de K funcții cunoscute cu coeficienți necunoscuți de ecuație:

$$m(p) = \sum_{i=1}^K a_i \cdot f_i(p) . \quad (2.73)$$

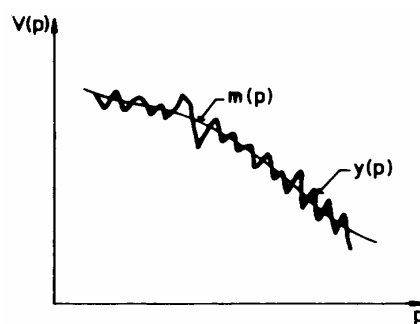


Fig. 2.27 Modelul funcției aleatoare nestaționare

Cel mai frecvent, pentru a ține cont de prezența derivatei în vecinătatea de estimare este suficientă adoptarea unei forme polinomiale limitate la gradul unu (*derivă lineară*):

$$m(p) = a_1 + a_2 \cdot p \quad (2.74)$$

sau doi (*derivă pătratică*):

$$m(p) = a_1 + a_2 \cdot p + a_3 \cdot p^2 \quad (2.75)$$

Funcție de dimensiunea spațiului în care se face estimarea, deoarece p semnifică un punct din spațiu, forma tendinței este diferită:

- dacă spațiul este *unidimensional*, dimensiunea fiind spre exemplu axa timpului t , formele tendințelor *liniare* și *pătratice* sunt:

$$\begin{cases} m(t) = a_1 + a_2 \cdot t \\ m(t) = a_1 + a_2 \cdot t + a_3 \cdot t^2 \end{cases} \quad (2.76)$$

- dacă spațiul este *bidimensional*, reprezentat într-un sistem de referință de coordonate rectangulare x și y , formele celor două tendințe sunt:

$$\begin{cases} m(x, y) = a_1 + a_2 \cdot x + a_3 \cdot y \\ m(x, y) = a_1 + a_2 \cdot x + a_3 \cdot y + a_4 \cdot x^2 + a_5 \cdot x \cdot y + a_6 \cdot y^2 \end{cases} \quad (2.77)$$

- dacă spațiul este *tridimensional*, de coordonate x , y și z forma tendințelor este:

$$\begin{cases} m(x, y, z) = a_1 + a_2 \cdot x + a_3 \cdot y + a_4 \cdot z \\ m(x, y, z) = a_1 + a_2 \cdot x + a_3 \cdot y + a_4 \cdot z + a_5 \cdot x^2 + a_6 \cdot y^2 + a_7 \cdot z^2 + a_8 \cdot x \cdot y + a_9 \cdot x \cdot z + a_{10} \cdot y \cdot z \end{cases} \quad (2.78)$$

În cazul spațiului tridimensional de obicei tendința este mai sensibilă într-o anumită direcție astfel încât forma ei analitică se simplifică. Dacă tendința se manifestă numai pe direcția verticală (z), variabila fiind staționară în plan orizontal (xOy), tendința pătratică se reduce la forma:

$$m(x, y, z) = a_1 + a_2 \cdot z + a_3 \cdot z^2 \quad (2.79)$$

Covarianța și variograma funcției aleatoare nestaționare. Pentru funcția aleatoare nestaționară cu structura din ecuația (2.72) formulele covarianței și variogramei sunt:

$$C(p_1, p_2) = E\{V(p_1) \cdot V(p_2)\} - m(p_1) \cdot m(p_2) = E\{Y(p_1) \cdot Y(p_2)\} \quad (2.80)$$

și:

$$\begin{aligned} 2\gamma(p_1, p_2) &= Var\{V(p_1) - V(p_2)\} = E\{[V(p_1) - V(p_2)]^2\} - \\ &- [m(p_1) - m(p_2)]^2 = Var\{Y(p_1) - Y(p_2)\} = E\{[Y(p_1) - Y(p_2)]^2\} \end{aligned} \quad (2.81)$$

Variograma din formula (2.81), adică aceea a rezidului *real*, nu poate fi estimată pornind de la datele originale în cazul prezenței unei tendințe. Pentru calculul variogramei adevărate ar trebui estimate simultan *deriva* și *variograma*, plecând de la un singur set de date, problemă care nu are o soluție unică riguroasă.

O metodă aproximativă pentru inferența simultană a tendinței și variogramei impune parcurgerea următoarelor etape de prelucrare:

- alegerea unui *model de variogramă*, de cele mai multe ori acesta fiind linear și izotrop;
- estimarea tendinței în fiecare punct de observație ($m(p_i)$) pe baza modelului de variogramă ales;
- calculul variogramei reziduurilor experimentale;
- compararea erorilor introduse de modelul de variogramă ales cu cele introduse de variograma calculată pe baza reziduurilor experimentale;
- adoptarea modelului de variogramă ales (în cazul concordanței erorilor introduse de cele două variograme) sau alegerea unui alt model de variogramă și reluarea prelucrării de la prima etapă.

Experiența arată că în cea mai mare parte a cazurilor se poate adopta fie o variogramă cvasistaționară determinată pe zone vecine ale zonei de estimare cu o corecție de plafon fie o variogramă lineară calculată pe baza comportării variogramei în vecinătatea originii.

Necunoașterea covarianței sau variogramei adevărate face ca prin kriging universal să nu se poată atinge valoarea minimă a varianței erorii de estimare. Acest lucru poate fi neglijat uneori deoarece în cazul prezenței tendinței nu ne interesează determinarea tendinței ci minimizarea incertitudinii estimării datorată acesteia.

Pe lângă tehnicile iterative utilizate la determinarea variogramei și covarianței *adevărate*, pentru stabilirea sistemului de kriging universal se apelează la *covarianța generalizată* a cărei inferență este posibilă pornind de la un set unic de date (P. Delfiner & Matheron, 1980;).

Noțiunea de *covarianță generalizată* este legată de funcția aleatoare intrinsecă de ordinul k , o generalizare pentru funcția aleatoare staționară corespunzătoare ordinului $k = 0$.

Trecerea de la funcția *aleatoare staționară* utilizată în cadrul kriging-ului punctual ordinar la funcția *aleatoare intrinsecă de ordinul zero* se face prin:

- înlocuirea covarianței $c(h)$ prin variogramă $\gamma(h)$. Se câștigă în acest mod în generalitate, clasa *variogramelor* fiind mult mai extinsă decât a *covarianțelor*. Variograma, nefiind limitată, permite descrierea variabilelor cu o dispersie teoretic nelimitată. Astfel, suprafețele piezometrice admit o variogramă lineară dar nu au covarianță staționară;
- utilizarea variogramei permite studiul variabilelor care nu au o speranță matematică constantă prin analiza creșterilor variabilei.

Ecuatiile sistemului pentru kriging universal. În cazul kriging-ului universal, estimatorul variabilei într-un punct p_0 este dat de expresia lineară:

$$V_0^* = \sum_{i=1}^n w_i \cdot v_i . \quad (2.82)$$

Condițiile pe care trebuie să le respecte estimarea sunt aceleași ca și în cazul kriging-ului ordinar (*linearitate, estimare nedeviată și minimizarea varianței erorii de estimare*), adăugându-se condiții suplimentare datorate prezenței tendinței.

În cazul prezenței unei tendințe de forma (2.73) condiția de estimare nedeviată devine:

$$\sum_{i=1}^n w_i \sum_{l=1}^k a_l \cdot f_l(p_i) - \sum_{l=1}^k a_l \cdot f_l(p_0) = 0 , \quad (2.83)$$

din care, deoarece coeficienții derivatei sunt necunoscuți, trebuie ca:

$$\sum_{i=1}^n w_i \cdot f_l(p_i) = f_l(p_0), \quad (2.84)$$

pentru l luând valori de la 1 la K , K fiind gradul maxim al polinomului ce modelează tendința.

Variograma este legată de varianța de estimare prin relația:

$$\tilde{\sigma}_R^2 = 2 \sum_{i=1}^n w_i \cdot \tilde{\gamma}_{i0} - \sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{\gamma}_{ij} \quad (2.85)$$

astfel încât pentru minimizarea ei în condițiile unei estimări nedeviate *sistemul de kriging universal* este:

$$\begin{cases} \sum_{j=1}^n w_j \cdot \tilde{\gamma}_{ij} + \sum_{l=1}^k \mu_l \cdot f_l(p_i) = \gamma_{i0}, i = 1, 2, \dots, n \\ \sum_{i=1}^n w_i \cdot f_l(p_i) = f_l(p_0), l = 1, 2, \dots, k \end{cases}, \quad (2.86)$$

sau sub formă matricială, în cazul unei tendințe de forma (2.77) pentru un spațiu bidimensional:

$$\begin{bmatrix} \tilde{\gamma}_{11} & \tilde{\gamma}_{12} & \cdots & \tilde{\gamma}_{1n} & 1 & x_1 & y_1 \\ \tilde{\gamma}_{21} & \tilde{\gamma}_{22} & \cdots & \tilde{\gamma}_{2n} & 1 & x_2 & y_2 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ \tilde{\gamma}_{n1} & \tilde{\gamma}_{n2} & \cdots & \tilde{\gamma}_{nn} & 1 & x_n & y_n \\ 1 & 1 & \cdots & 1 & 0 & 0 & 0 \\ x_1 & x_2 & \cdots & x_n & 0 & 0 & 0 \\ y_1 & y_2 & \cdots & y_n & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} \tilde{\gamma}_{10} \\ \tilde{\gamma}_{20} \\ \vdots \\ \tilde{\gamma}_{n0} \\ 1 \\ x_0 \\ y_0 \end{bmatrix} \quad (2.87)$$

Relația de calcul pentru varianța erorii de estimare este:

$$\tilde{\sigma}_R^2 = \sum_{i=1}^n w_i \cdot \tilde{\gamma}_{i0} + \sum_{j=1}^k \mu_j \cdot f_j(p_0). \quad (2.88)$$

De reținut că pentru realizarea kriging-ului în prezența unei tendințe regionale nu este necesar calculul coeficienților funcției care o modelează ci numai forma ei. Dacă forma tendinței este prost aleasă varianța erorilor de estimare va fi mare chiar dacă termenii de ordin superior ai derivatei au valori mici și dacă intervalele de încredere ale valorilor interpolate sunt în realitate mici.

Eficiența programului pentru kriging-ul universal crește dacă este prevăzută posibilitatea utilizării unor tendințe de forme cât mai complexe, chiar dacă timpul de calcul este în acest fel prelungit.

A23. Evaluarea distribuției unei variabile nestaționare

Să se estimeze distribuția spațială a sarcinii piezometrice a acviferului cu nivel liber din conul aluvionar Ialomița-Dâmbovița, în vecinătatea captării de apă subterană a orașului Târgoviște (Fig.2.28), pe baza cotelor nivelurilor piezometrice măsurate în forajele hidrogeologice de explorare și captare.

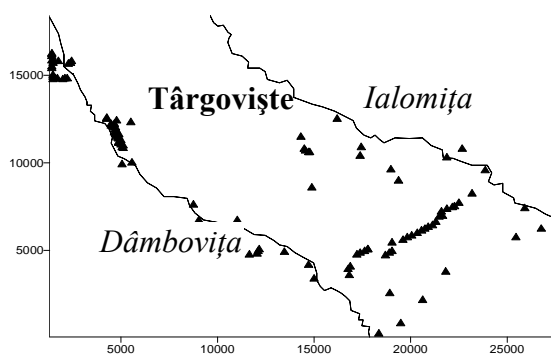


Fig.2.28. Distribuția celor 113 foraje hidrogeologice de explorare și captare

Rezolvare:

Sarcina piezometrică a acviferelor este o variabilă nestaționară tipică. Analiza distribuției în spațiu a sarcinii piezometrice permite identificarea direcțiilor de curgere ale apelor subterane și stă la baza amplasării și dimensionării captărilor de apă subterană.

Pentru estimarea corectă a distribuției sarcinii piezometrice trebuie identificate:

- direcțiile de anizotropie;
- forma tendinței ce afectează staționaritatea structurii.

Direcțiile de anizotropie ale distribuției spațiale a sarcinii piezometrice pentru zona studiată se determină pe baza hărții conturale a variogrammei de suprafață (Fig.2.29) și sunt:

- N45°E pentru direcția de continuitate maximă (↗ în Fig.2.29);
- N45°V pentru direcția de continuitate minimă.

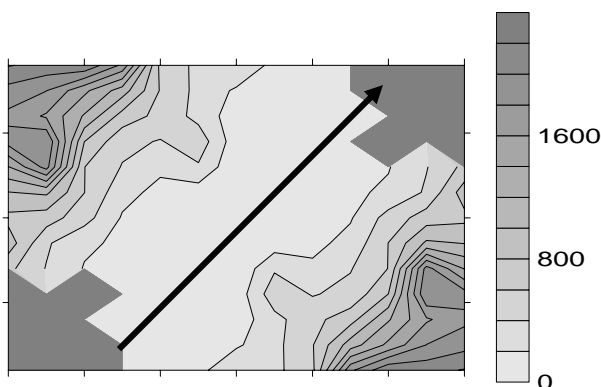


Fig.2.29. Variograma de suprafață a cotei nivelului piezometric.

Modelul variogrammei pe aceste două direcții este de tip linear cu un raport de anizotropie $r_a = 12$.

Raportul de anizotropie s-a calculat ca raport al razelor de influență ale variogramelor direcționale, la valoarea variogrammei experimentale de 150 m^2 , pentru principalele direcții de anizotropie:

- pe direcția N45°E: $r_1 = 12000 \text{ m}$;
- pe direcția N45°V: $r_2 = 1000 \text{ m}$

Variogramele direcționale au efectul de pepită nul, eliminându-se astfel orice suspiciune asupra existenței unor structuri “îngropate” ce nu au putut fi identificate datorită distanței prea mari dintre forajele în care s-au măsurat nivelurile piezometrice. Acest lucru este argumentat și de caracterul compact al histogramei nivelurilor piezometrice care indică o selecție omogenă de valori (**Fig.2.30**).

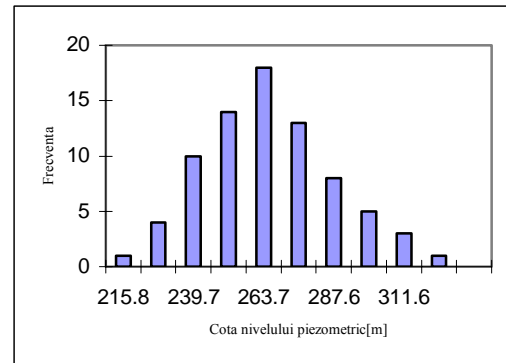


Fig.2.30. Histograma cotelor nivelului piezometric.

Distribuția valorilor măsurate ale nivelului piezometric indică o tendință generală de creștere a cotelor acestuia de la SE spre NV pe direcția $N45^{\circ}V$. Pentru modelarea acestei tendințe s-au adoptat două forme analitice: *lineară* (2.74) și *pătratică* (2.75).

Compararea rezultatelor obținute în cele două variante (**Fig.2.31** și **Fig.2.32**) recomandă ca *optim modelul pătratic* al tendinței de variație a cotei sarcinii piezometrice.

Certificarea caracterului optim al modelului pătratic pentru tendința zonală a distribuției sarcinii piezometrice se face în etapa finală de calcul al erorilor de estimare a distribuției spațiale.

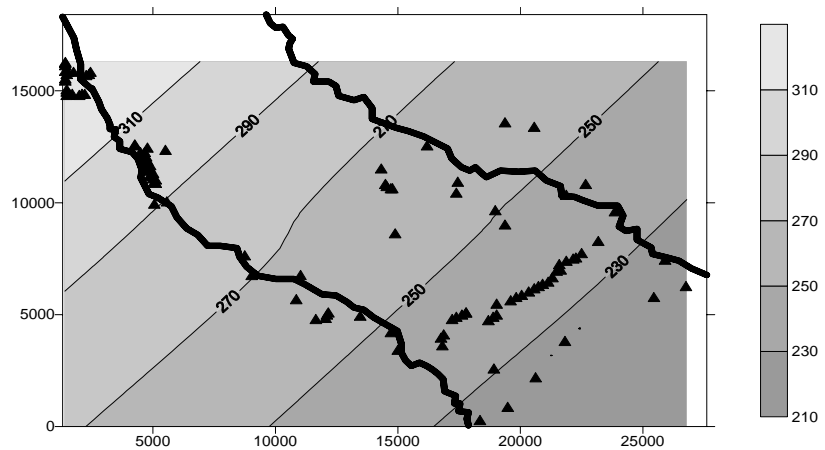


Fig.2.31. Estimarea distribuției sarcinii piezometrice cu model liniar al tendinței

Neglijarea caracterului nestăionar al distribuției cotei nivelului piezometric conduce la erori de estimare a distribuției acesteia de ± 20 m în partea SE-ică a zonei studiate (**Fig.2.33**).

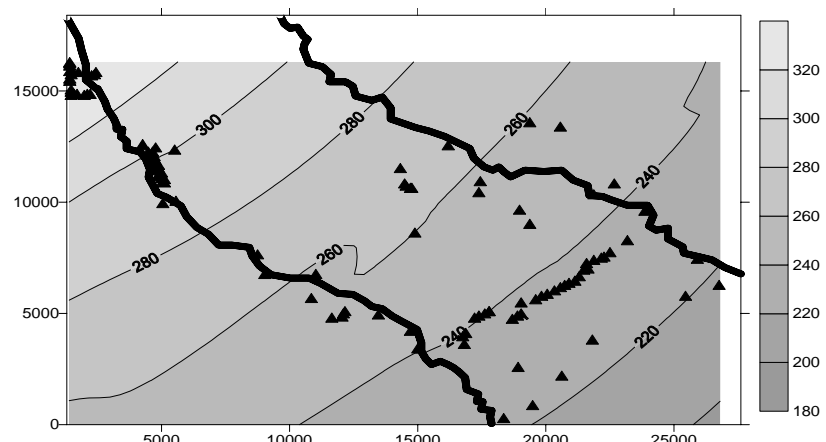


Fig.2.32. Estimarea distribuției sarcinii piezometrice cu model pătratic al tendinței

Erorile de estimare s-au calculat prin compararea estimărilor realizate fără eliminarea componentei de tendință a cotei sarcinii piezometrice și prin eliminarea acesteia pe baza modelului pătratic (2.75).

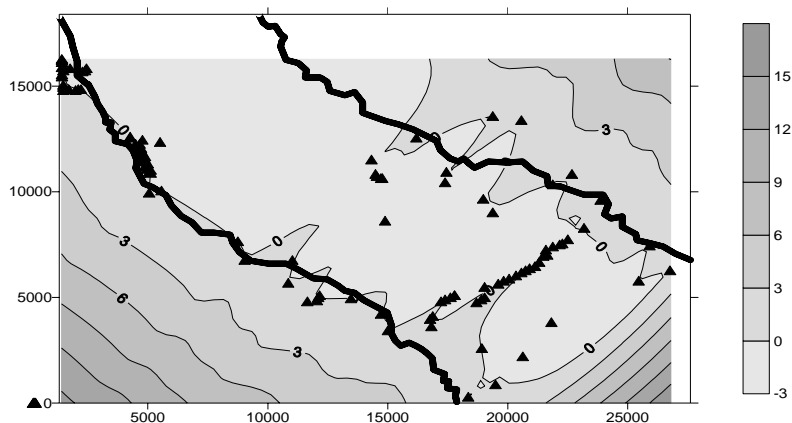


Fig.2.33. Distribuția erorilor de estimare datorate neglijării tendinței pătratice a cotei sarcinii piezometrice

COMENTARIU

Neglijarea caracterului nestaționar al structurilor spațiale conduce la erori semnificative în estimarea acestora.

Primul moment dificil al estimării este sesizarea caracterului nestaționar !

Să ne reamintim că prezența caracterului nestaționar se face simțită în forma variogramelor experimentale prin lipsa stabilizării acestora în jurul unei valori constante (absența palierului).

Sesizarea nestaționarității este posibilă și prin construirea unor grafice care să illustreze variația parametrului măsurat de-a lungul unor profile rectilinii. În astfel de reprezentări oscilațiile valorilor parametrului studiat în jurul unei valori constante indică staționaritatea acestuia iar creșterea sau descreșterea continuă indică nestaționaritatea.

Al doilea moment dificil este alegerea modelului analitic al tendinței !

Alegerea modelului optim al tendințelor nestaționare se face prin compararea erorilor introduse de două sau mai multe modele calibrate pe datele disponibile. Abaterile de la datele măsurate, introduse de modele, se compară la nivel global pe baza parametrilor statistici (medie, dispersie, coeficient de variație) înainte de estimarea distribuției spațiale.

Distribuția spațială a erorilor de estimare introduse de neadecvarea modelului tendinței este calculată simultan cu estimarea distribuției spațiale în care modelul respectiv este implicat. Dacă modelul a fost prost ales rezultă erori de estimare mari și se revine la etapa eliminării nestaționarității alegând alt model pentru tendință.

Datorită variabilității complexe a caracteristicilor studiate distribuția erorilor nu va fi niciodată uniformă. În funcție de modelul ales, în anumite zone vor fi erori mai mari iar în altele mai mici.

De cele mai multe ori un model unic al tendinței nu va conduce la erori minime în toată zona studiată!

2.3. ESTIMAREA ZONALĂ

Obiectivul acestui capitol este prezentarea metodologiei de prognoză a *valorii medii* a unei variabile *pe suprafețe* de forme și extinderi variabile.

Această metodologie servește printre altele: *evaluării rezervelor* de substanțe minerale utile solide, *inițializării parametrice a modelelor de simulare* numerică a proceselor potențiale (exemplu: câmpul de gravitație, câmpul termic, câmpul vitezelor de curgere a unui fluid etc.)

Kriging-ul punctual este o metodă performantă și pentru estimarea valorii medii a unei variabile într-un domeniu spațial bidimensional sau tridimensional limitat de un contur oarecare.

Fără a introduce o modificare operațională esențială *kriging-ul punctual* permite calculul valorii medii prin discretizarea domeniului spațial și medierea valorilor estimate în punctele de discretizare. Deși conceptual simplă, această metodă devine costisitoare prin volumul mare de calcule pe care îl implică.

Pentru reducerea volumului de calcule, fără reducerea performanței estimării, *kriging-ul zonal* operează numai modificarea matricii D a sistemului de *kriging punctual* (2.34).

Ecuatiile sistemului de kriging zonal. Realizând estimarea în aceleași condiții cu ale *kriging-ului punctual ordinar* și *universal* (estimare lineară nedeviată cu varianță minimă a erorilor de estimare) *kriging-ul zonal* presupune rezolvarea unui sistem similar cu cel din ecuația (2.33) scris în raport cu modelul de covarianță, în care se modifică doar matricea D .

Matricea C constituită din valorile covarianței variabilei v , calculată între punctele de observație, este în mod evident independentă de punctul sau zona în care se face estimarea. Ea va rămâne nemodificată în cazul sistemului de *kriging zonal* având același rol de eliminare a efectului grupării neregulate a punctelor de observație.

Matricea D este constituită din covarianțele calculate pe baza valorilor variabilelor aleatoare de pe poziția *punctelor de probare* (p_1, p_2, \dots, p_9 ; **Fig.2.34**) și a valorilor din *punctele de estimare* ($A_1, A_2, A_3, A_4, \dots, A_6$; **Fig.2.34**). Pentru estimarea punctuală aceste covarianțe sunt calculate doar între două puncte iar pentru estimarea zonală covarianțele se calculează între punctele de măsură (p_1, p_2, \dots, p_9) și zona pe care se estimează valoarea medie (suprafața A hașurată; **Fig.2.34**).

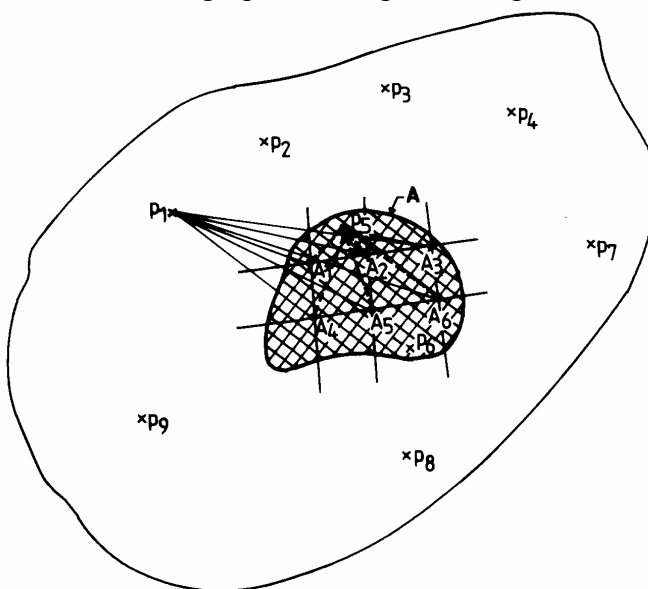


Fig. 2.34. Kriging zonal

Covarianța punct-zonă (cov_{iA}) este evaluată tot pe baza unui model de covarianță calat pe covarianța experimentală, cu relația:

$$\begin{aligned}
 \tilde{c}_{iA} &= Cov\{V_A V_i\} = E\{V_A V_i\} - E\{V_A\} \cdot E\{V_i\} = E\left\{\frac{1}{n_A} \sum_{j \in A} V_j V_i\right\} - \\
 &- E\left\{\frac{1}{n_A} \sum_{j \in A} V_j\right\} E\{V_i\} = \frac{1}{n_A} \sum_{j \in A} E\{V_j V_i\} - \frac{1}{n_A} \sum_{j \in A} E\{V_j\} \cdot E\{V_i\} = \\
 &= \frac{1}{n_A} \sum_{j \in A} Cov\{V_j V_i\}
 \end{aligned} \tag{2.89}$$

Covarianța între variabila aleatoare de pe poziția p_i (V_i) și cea reprezentând valoarea medie a variabilei pe suprafața A este conform ecuației (2.89) o medie a covarianțelor dintre variabila V_i și cele din toate punctele de discretizarea (n_A) ale suprafeței A (V_{A1}, \dots).

Sistemul de kriging zonal, scris în funcție de covarianță și obținut prin minimizarea varianței erorii de estimare, este:

$$\begin{bmatrix} \tilde{c}_{11} & \tilde{c}_{12} & \dots & \tilde{c}_{1n} & 1 \\ \tilde{c}_{21} & \tilde{c}_{22} & \dots & \tilde{c}_{2n} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \tilde{c}_{n1} & \tilde{c}_{n2} & \dots & \tilde{c}_{nn} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \bar{w}_1 \\ \bar{w}_2 \\ \vdots \\ \bar{w}_n \\ \mu_A \end{bmatrix} = \begin{bmatrix} \frac{1}{n_A} (\tilde{c}_{1A1} + \tilde{c}_{1A2} + \dots + \tilde{c}_{1An_A}) \\ \frac{1}{n_A} (\tilde{c}_{2A1} + \tilde{c}_{2A2} + \dots + \tilde{c}_{2An_A}) \\ \vdots \\ \frac{1}{n_A} (\tilde{c}_{nA1} + \tilde{c}_{nA2} + \dots + \tilde{c}_{nAn_A}) \\ 1 \end{bmatrix} = \begin{bmatrix} \bar{c}_{1A} \\ \bar{c}_{2A} \\ \vdots \\ \bar{c}_{nA} \\ 1 \end{bmatrix} \tag{2.90}$$

Soluția sistemului de kriging zonal constă în valorile ponderilor medii \bar{w}_i acordate fiecărei valori măsurate pentru calculul valorii medii pe suprafața A cu relația:

$$v_A^* = \sum_{i=1}^n \bar{w}_i \cdot v_i \tag{2.91}$$

și a parametrului lui Lagrange care minimizează varianța erorii de estimare calculată cu relația:

$$\tilde{\sigma}_A^2 = \tilde{c}_{AA} - \left(\sum_{i=1}^n \bar{w}_i \cdot \tilde{c}_{iA} + \mu_A \right) \tag{2.92}$$

în care \tilde{c}_{AA} este covarianța medie a suprafeței A obținută cu relația:

$$\tilde{c}_{AA} = \frac{1}{n_A^2} \sum_{i \in A} \sum_{j \in A} \tilde{c}_{ij} \tag{2.93}$$

Valoarea medie a covarianței suprafeței A se calculează prin discretizarea suprafeței A prin aceleași puncte utilizate la calculul covarianțelor componente ale matricii D ($c_{1A}, c_{2A}, \dots, c_{nA}$).

Precizia kriging-ului zonal. Precizia evaluării valorii medii pentru o suprafață aleasă (A) este determinată de numărul de puncte de discretizare (n_A) și geometria amplasării acestora.

Numărul punctelor de discretizare ($A_i, i = 1, 2, \dots, n_A$) este proporțional cu precizia de estimare a kriging-ului zonal. Densitatea optimă a punctelor de discretizare se obține experimental și este cea de la care valoarea estimată se stabilizează. Experiența arată că numărul maxim de puncte de discretizare necesare pe unitatea de suprafață este de 16 (Fig.2.35).

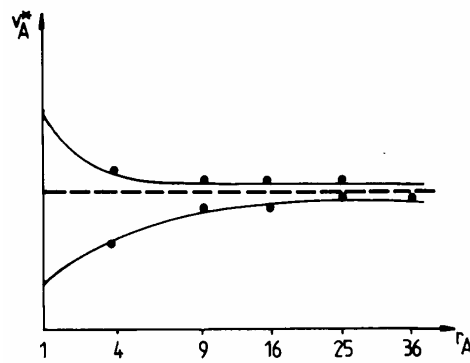


Fig.2.35. Influența numărului de puncte de discretizare în kriging-ul zonal

Geometria optimă a punctelor de discretizare este determinată de caracterul izotrop sau anizotrop al structurii spațiale.

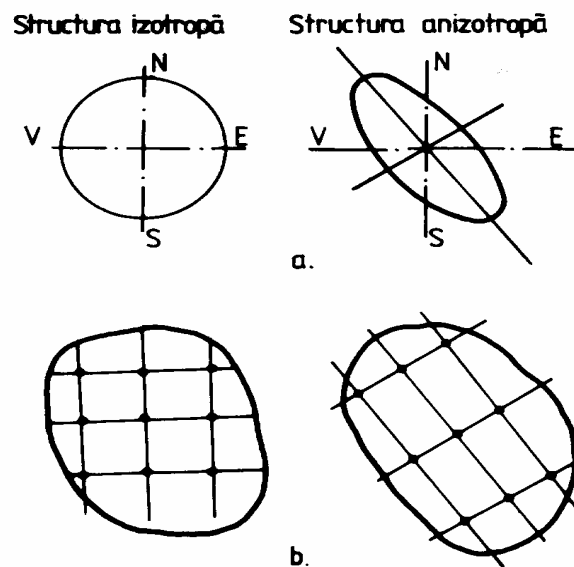


Fig. 2.36. Rețeaua de discretizare pentru kriging zonal: a - diagramele radiare ale variogramelor; b - geometria punctelor de discretizare

numărului de puncte de discretizare și a amplasării lor presupune experiența de prelucrare și un soft interactiv.

Particularități operaționale. Avantajul utilizării kriging-ului zonal este cel al obținerii valorii medii pe suprafața A din rezolvarea unui singur sistem (2.90). Complicația introdusă de utilizarea sistemului (2.90) în raport cu cea a sistemului (2.34) este doar cea a calculului covarianțelor medii (2.89). Acest calcul suplimentar este mai puțin consumator de timp în raport cu rezolvarea unui număr de sisteme de tip (2.34) egal cu numărul punctelor de discretizare.

Posibilitatea utilizării kriging-ului punctual și zonal pentru obținerea aceleiași valori medii pentru o suprafață oarecare se bazează pe următoarele proprietăți rezultate din caracterul linear al metodei de prelucrare:

Pentru structurile izotrope (Fig.2.36a) rețeaua de discretizare nu are o orientare preferențială și pentru eficiența prelucrării ea se orientează paralel cu axele de coordonate ale sistemului în care se amplasează punctele de observație. Rețeaua de discretizare este pătratică în acest caz (Fig.2.36b).

Pentru structurile anizotrope (Fig.2.36a) rețeaua de discretizare se orientează paralel cu direcțiile de anizotropie rezultate din diagrama radiară a variogramei, iar distanța dintre punctele de discretizare este mai mică pe direcția de continuitate minimă și mai mare pe direcția de continuitate maximă (Fig.2.36b).

Controlul preciziei de estimare a kriging-ului zonal prin alegerea

- media valorilor estimate prin kriging punctual (utilizând sistemul (2.34)) în cele n_A puncte de discretizare este egală cu valoarea medie obținută prin kriging zonal (utilizând sistemul 2.90) pe baza aceluiași puncte de discretizare;
- media ponderilor (w_i) acordate unui punct de probare în raport cu punctele de discretizare este egală cu ponderea zonală a punctului de probare în raport cu întreaga zonă evaluată (\bar{w}_i).

Compatibilitatea evaluării zonale și punctuale este valabilă numai pentru kriging, alte metode nu pot fi adaptate în aceeași manieră. Spre exemplu, utilizarea metodei inversului distanței în aceeași manieră pentru estimările zonale nu conduce la aceleași rezultate.

În **Fig.2.37**, pe suprafața A sunt amplasate două puncte de discretizare A_1 și A_2 utilizate pentru estimarea valorii medii v_A , pe baza valorilor măsurate v_1 , v_2 , v_3 și v_4 .

Calculul valorilor punctuale în A_i ($i = 1,2$) se face cu relația:

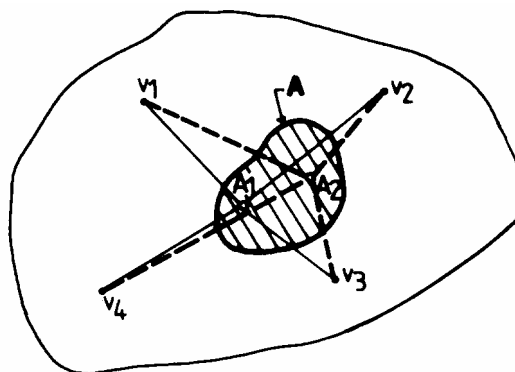


Fig. 2.37.Metoda inversului distanței pentru evaluare punctuală și zonală.

$$v_{A_i}^* = \frac{\frac{v_1}{d_{A_1 v_1}^p} + \frac{v_2}{d_{A_1 v_2}^p} + \frac{v_3}{d_{A_1 v_3}^p} + \frac{v_4}{d_{A_1 v_4}^p}}{\frac{1}{d_{A_1 v_1}^p} + \frac{1}{d_{A_1 v_2}^p} + \frac{1}{d_{A_1 v_3}^p} + \frac{1}{d_{A_1 v_4}^p}}, \quad (2.94)$$

în care:

v_i - valorile măsurate ($i = 1, 2, 3, 4$);

$d_{A_i v_i}$ - distanța dintre punctul A_i și punctul în care se măsoară v_i ;

p - număr real pozitiv care frecvent are valoarea 2 în metoda inversului distanței (dacă valorile lui p sunt subunitare, ponderile acordate valorilor măsurate se egalizează iar dacă p are valori supraunitare ponderile se diferențiază, crescând cea a valorilor din vecinătatea punctului de estimare).

Dacă se estimează valoarea medie pe suprafața A ca o medie aritmetică a valorilor estimate în cele două puncte de discretizare (A_1 și A_2), aceasta va fi diferită de cea calculată cu relația:

$$v_A^* = \frac{\frac{v_1}{d_{v_1 A}^p} + \frac{v_2}{d_{v_2 A}^p} + \frac{v_3}{d_{v_3 A}^p} + \frac{v_4}{d_{v_4 A}^p}}{\frac{1}{d_{v_1 A}^p} + \frac{1}{d_{v_2 A}^p} + \frac{1}{d_{v_3 A}^p} + \frac{1}{d_{v_4 A}^p}}, \quad (2.95)$$

în care

$d_{v_i A}$ este distanța medie dintre valoarea v_i și suprafața A calculată cu relația:

$$d_{v_i A} = \frac{1}{2}(d_{A_1 v_i} + d_{A_2 v_i}); \quad i=1, 2, 3, 4. \quad (2.96)$$

Este evident că :

$$\frac{1}{2}(v_{A_1}^* + v_{A_2}^*) \neq v_A^* \quad (2.97)$$

și deci metoda inversului distanței nu are proprietățile kriging-ului în raport cu estimările punctuale și zonale.

Un exemplu de kriging zonal. Estimarea zonală prin kriging este exemplificată pentru evaluarea grosimii medii a acviferului sub presiune interceptat de șase foraje (**Fig.2.38**).

Estimarea valorii medii pe suprafața *A* (**Fig.2.38**) este realizată pe baza a patru puncte de discretizare (*A₁*, *A₂*, *A₃*, *A₄*) în două variante:

- media valorilor estimate prin *kriging punctual* în cele patru puncte de discretizare (prin rezolvarea a patru sisteme de kriging punctual de forma (2.33));
- estimarea prin *kriging zonal* (prin rezolvarea unui singur sistem de kriging zonal de forma (2.90)).

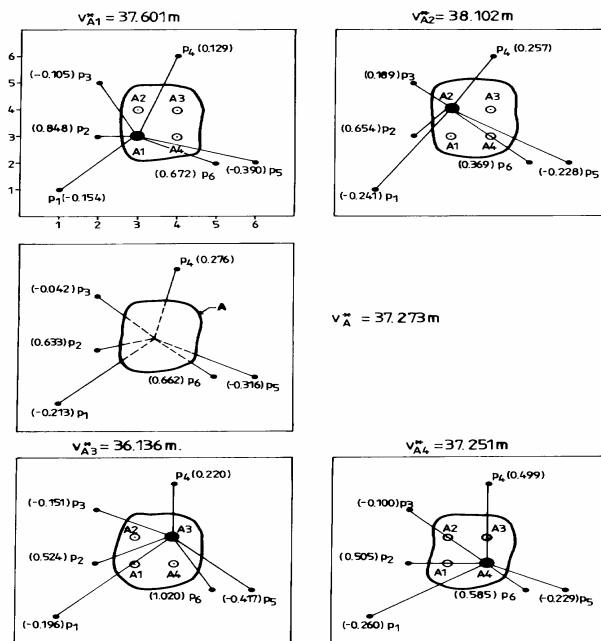


Fig. 2.38. Calculul valorii medii pe suprafața *A* prin kriging zonal.

Tabelul 2.10 Ponderile kriging-ului punctual și zonal

Pct. Obs.	Ponderile					<i>v_i</i>
	kriging punctual (<i>w_i</i>)				kriging zonal (<i>w_i</i>)	
	A1	A2	A3	A4		
P ₁	-0,154	-0,241	-0,196	-0,260	-0,213	25
P ₂	0,848	0,654	0,524	0,505	0,633	40
P ₃	-0,105	0,189	-0,151	-0,100	-0,042	38
P ₄	0,129	0,257	0,220	0,499	0,276	32
P ₅	-0,390	-0,228	-0,417	-0,229	-0,316	52
P ₆	0,672	0,369	1,020	0,585	0,662	40

În **tabelul 2.10** sunt sintetizate ponderile (*w_i*) acordate valorilor măsurate în cele două variante: *kriging punctual* și *kriging zonal*. Modelul de variogramă utilizat este *izotrop de tip gaussian* cu *palierul c_o = 10 m²* și *raza r = 6 km*.

Fig.2.38 prezintă poziția punctelor de observație și ponderile asociate fiecăruia în raport cu punctele de discretizare. Ponderile, scrise în paranteze, precum și valorile estimate verifică proprietățile particulare ale *kriging-ului zonal*:

- *media ponderilor* acordate valorii v_i măsurată în punctul p_i ($i = 1,2,\dots,6$) în *kriging-ul punctual* pentru calculul valorilor din A_1, A_2, A_3, A_4 este egală cu ponderea acordată în raport cu suprafața A în *kriging-ul zonal*. Pentru $i = 3$, adică pentru punctul p_3 , se obține:

$$\bar{p}_3 = (-0,105 + 0,189 - 0,151 - 0,100)/4 = -0,042$$

- *media valorilor calculate prin kriging punctual* în punctele A_1, A_2, A_3, A_4 este egală cu valoarea calculată prin *kriging zonal*:

$$\bar{v}_A^* = (37,601 + 38,102 + 36,136 + 37,521)/4 = 37,273$$

A24. Calcul de rezerve

Să se calculeze rezerva de plumb dintr-un paralelipiped dreptunghic având baza un pătrat cu latura de 100 m și înălțimea de 1 m amplasat în NE-ul zonei probate din zăcământul Dealul Roatei (în **Fig.2.43** este reprezentat printr-un pătrat conturat cu linie groasă întreruptă).

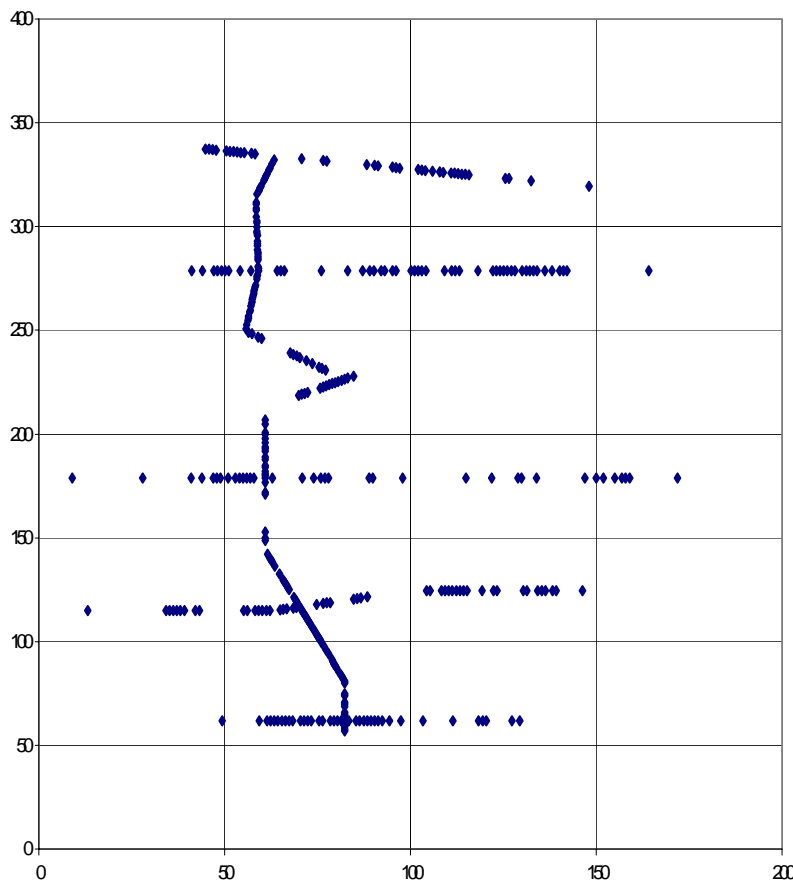


Fig.2.39. Rețeaua de probarea a zăcamântului Dealul Roatei

Rezolvare:

Metoda *kriging-ului zonal* este cea care permite evaluarea *cele mai probabile valori medii* a conținuturilor pe o *suprafață oarecare*, prin eliminarea efectului negativ al distribuției neregulate a punctelor de probare și a variabilității mari a conținuturilor.

Zăcământul Dealul Roatei, din punct de vedere geosuctural, se încadrează în unitatea Munților Apuseni de Sud. Mineralizația este cantonată în formațiunea conglomeratului de Fața

Băii, de vârstă paleogenă. Longitudinal, zăcământul a fost urmărit pe două galerii de aproximativ 350 m. Probarea care stă la baza exemplificării calculului de rezerve s-a realizat la nivelul orizontului +516 m pe galeriile transversale (**Fig.2.39**). Zăcământul a fost descoperit și cercetat de MINEXFOR DEVA.

S-a recoltat un număr de 1002 probe din care s-au determinat conținuturile pentru Au, Pb, Zn, Cu, Hg, Tl, Ga, In, Ni, Bi, Mo, V, Cd, Co, Ba, Sr, Se, Sb.

În mod clasic, pentru calculul rezervei se determină o valoare medie a conținutului care se multiplică apoi cu volumul luat în considerare. Valoarea medie a conținutului este afectată de erori mai mari sau mai mici în funcție de numărul de probe utilizate, de distribuția acestora și de metoda de calcul utilizată. Kriging-ul zonal asigură eroarea minimă de evaluare a conținutului mediu pentru un anumit număr de probe.

Analiza variabilității globale a conținuturilor în plumb

indică o selecție de valori eterogenă cu distribuție lognormală.

Diagrama de probabilitate a celor 1002 conținuturi în plumb (**Fig.2.40**), prin forma curbilinie, evidențiază distribuția “anormală” a valorilor (în cazul unei distribuții normale toate cele 1002 puncte din diagramă ar fi trebuit să se dispună de-a lungul unei drepte).

Anamorfoza gaussiană a selecției de valori (realizată prin logaritmarele acestora) evidențiază trei grupuri de valori distincte din punct de vedere statistic (**Fig.2.41**):

- grupul G1, format din 257 de probe cu conținutul în plumb cuprins în intervalul [0% și 0,01%];
- grupul G2, format din 732 de probe cu conținutul în plumb cuprins în intervalul (0,01 și 1,00 %];
- grupul G3, format din 13 de probe cu conținutul în plumb cuprins în intervalul (1,02 și 23,00 %].

Grupul G1 are o caracteristică specială fiind alcătuit din 66 de valori de 0,0% plumb și 191 de valori de 0,01%Pb. Cele 257 de valori sunt distribuite pe toată suprafața explorată (**Fig.2.42**) iar caracteristicile structurii acestei distribuții sunt sintetizate în variograma omnidirecțională de tip sferic ale cărei caracteristici sunt sintetizate în **tabelul 2.11**.

Grupul G2 are o distribuție lognormală cu o variabilitate mai pronunțată a structurii spațiale (varianța totală este 0,039, cu câteva ordine de mărime mai mare decât a grupului G1; rezultă că valorile mai mici de 0,01% ale conținuturilor în plumb vor fi estimate mult mai corect față de cele mai mari decât această valoare).

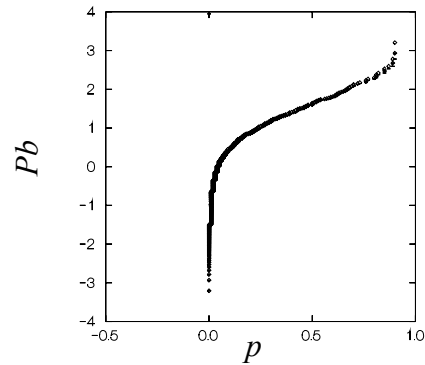


Fig.2.40.Diagrama de probabilitate a conținuturilor în Pb

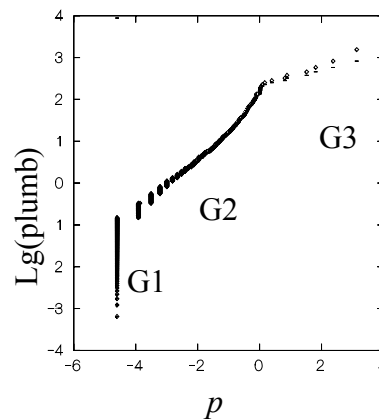


Fig.2.41.Diagrama de probabilitate (p) a conținuturilor în plumb

Tabelul 2.11. Caracteristicile variogramelor pentru G1 și G2.

Grup	Interval valoric	Model variograma	Efect de pepită	Varianță totală	Rază de influență
G1	[0,0-0,1%Pb]	Sferic	$1,54 \times 10^{-5}$	$2,16 \times 10^{-5}$	40
G2	(0,1-1,0%Pb]	Sferic	$2,70 \times 10^{-2}$	$3,90 \times 10^{-2}$	40

Grupul G3, numeric nesemnificativ în raport cu grupurile G1 și G2 (numai 13 valori, adică 1% din numărul total de probe prelucrate; G1 reprezintă numeric 26% din selecția de date iar G2: 73%) este exclus din evaluarea rezervei de plumb fără a afecta sensibil rezultatele estimărilor.

Pentru calculul rezervei de plumb din volumul ales prin kriging zonal se detaliază distribuția conținutului de plumb pe baza modelelor de variogramă determinate.

Calculul rezervelor de plumb pe baza metodei blocurilor geologice pentru volumul luat în considerare se face cu relația:

$$M = \sum_{i=1}^{i=n_x} \sum_{j=1}^{j=n_y} V_{ij} \cdot \gamma_{ij} \cdot c_{ij} \quad (2.98)$$

în care

V_{ij} - volumul blocului ij ($i = 1, \dots, n_x; j = 1, \dots, n_y$) [L^3];

n_x - numărul total de blocuri pe coloane;

n_y - numărul total de blocuri pe rânduri;

γ_{ij} - greutatea volumetrică a minereului din blocul ij [$MT^{-2}L^{-2}$];

c_{ij} - conținutul mediu al minereului [% (pentru Pb, Zn etc.) sau g/t (de obicei pentru Au, Ag și alte elemente cu conținuturi mici)].

În cazul studiat, calculul rezervei s-a realizat în 10 variante, pentru diferite valori ale lui n_x și n_y (tabelul 2.12).

Greutatea volumetrică a minereului a fost determinată de către ICITPLICIM Deva pe un număr redus de probe așa încât, în calculul rezervelor s-a luat o valoare unică pentru toate blocurile componente:

$$\gamma_{ij} = 2,60 \text{ tf/m}^3$$

Conținutul de plumb exprimat în procente a fost calculat pentru fiecare bloc de discretizare, în toate cele 10 variante, prin kriging, pe baza modelelor de variogramă indentificate pentru intervalele valorice corespunzătoare celor două grupuri (G1 și G2) (tabelul 2.11) și reprezentat grafic prin hărți cu izolinii:

- pentru domeniul valoric [0,0-0,1](G1, Fig.2.42a);
- pentru domeniul valoric (0,1-1,0)(G2, Fig.2.42b);
- pentru întreg domeniul valoric al selecției (realizată prin suprapunerea celor două hărți), hartă pe care este figurată și suprafața blocului pentru care se calculează rezerva de plumb (Fig.2.43).

Valoarea rezervei calculate variază în funcție de caracteristicile rețelei de discretizare (tabelul 2.12) din cauza erorilor introduse de metoda de interpolare (minime în cazul utilizării kriging-ului).

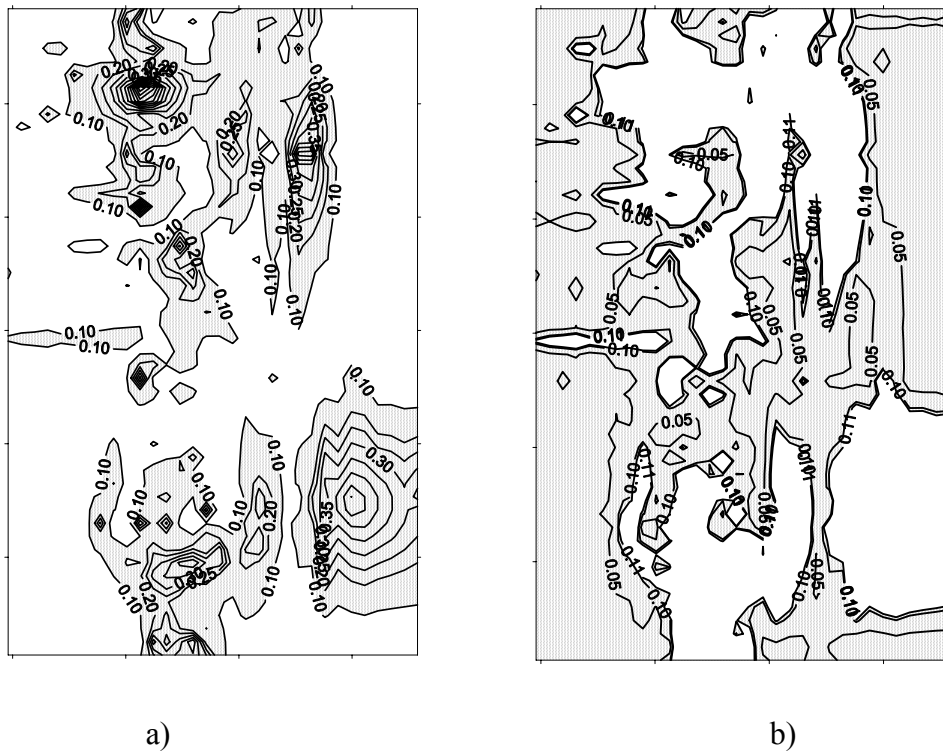


Fig.2.42. Distribuția conținutului de plumb pe intervale valorice ale grupurilor G1 și G2.

Tabelul 2.12. Valoarea rezervelor de Pb

Nr. Crt	n_x [-]	n_y [-]	γ [tf/m ³]	Rez. [t]
1	10	10	1,20	34.55
2	20	20	1,20	36.58
3	30	30	1,20	38.22
4	40	40	1,20	35.65
5	50	50	1,20	35.98
6	60	60	1,20	37.00
7	70	70	1,20	36.58
8	80	80	1,20	37.08
9	90	90	1,20	36.01
10	100	100	1,20	36.58

Valorile din **tabelul 2.12** pot fi calculate cu un program de tip *SURFER* care realizează atât operațiunea de interpolare de tip kriging, pentru calcul distribuției conținutului de plumb, cât și calculul rezervei.

Valoarea medie a rezervei calculată pe baza celor zece valori din **tabelul 2.12**, la un risc al erorii de gradul I $\alpha = 5\%$, este:

$$REZ(plumb) = 36,4 \pm 0,7 \text{ tone}$$

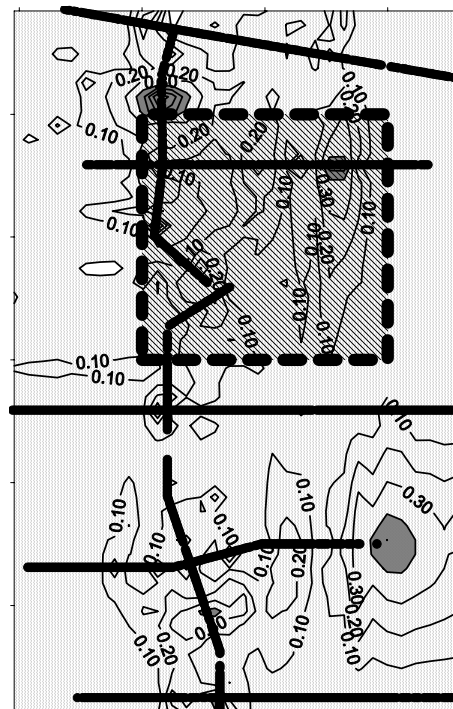


Fig.2.43. Distribuția conținutului de plumb și conturul blocului pentru calculul rezervei.

COMENTARIU

Problema centrală în calculul rezervelor de substanțe minerale utile este evaluarea corectă a conținutului mediu dintr-un volum de minereu excavat. Pe baza acestui conținut mediu se estimează dacă o acumulare de substanțe minerale utile este exploatabilă. Corectitudinea determinării conținutului mediu este premiza evaluării corecte a rezervelor unui zăcământ și a eficienței economice a exploatării acestuia.

Kriging-ul zonal este metoda actuală cea mai performantă pentru evaluarea rezervelor de substanțe minerale utile!

Pentru atingerea parametrilor optimi ai estimării rezervelor utilizând kriging-ul zonal, un rol determinant îl are separarea selecției de date disponibile în grupe omogene din punct de vedere statistic (ex.: G1 și G2). Neomogenitatea distribuției mineralizațiilor se poate manifesta sub două aspecte distincte:

- cel spațial, evidențiat prin variabilități diferite de la o zonă la alta a zăcământului (ex.: în nordul zăcământului să am un model al legității variației spațiale cuantificat printr-un model de variogramă sferică iar în sud printr-un model gaussian);*
- cel valoric, manifestat prin variabilități diferite pentru intervale valorice diferite (ex.: pentru valori ale conținutului în plumb mai mici de 0,1% modelul variabilității spațiale să fie o variogramă sferică iar pentru valori mai mari de 0,1% , modelul variabilității spațiale să fie o variogramă de tip linear).*

Kriging-ul zonal permite estimarea optimă a rezervelor prin reducerea variabilității caracteristicii studiate!

Eliminarea valorilor extreme și normalizarea distribuției selecțiilor de date conduc la estimarea cu ajutorul kriging-ului zonal a celei mai probabile rezerve. Această rezervă este însă calculată pe baza unei selecții de date finite care nu surprinde întotdeauna întreaga variabilitate a zăcămintelor.

Pentru protejarea investițiilor în domeniul exploatărilor miniere este necesar să fie evaluat riscul maxim pe care ni-l asumăm prin exploatarea unui zăcământ!

Soluționarea acestei probleme se face prin simularea condițională a distribuției spațiale a conținutului caracteristicii studiate. Simularea condițională încearcă să prognozeze variabilitatea maximă a zăcământului care poate fi "sugerată" de datele disponibile în limitele unei anumite legi de distribuție acceptate. Estimarea rezervelor se face pentru mai multe variante de variabilitate și se calculează un interval valoric în care acestea se pot plasa.

Prin kriging zonal se calculează rezerva minimizând eroarea de estimare iar prin simulare condițională se calculează rezerva în condițiile erorii maxime pe care o putem face pe utilizând selecția de date disponibile.

A25. Inițializarea parametrică a modelelor numerice

Să se inițializeze distribuția transmisivității acviferului cu nivel liber din interfluviul Ialomița-Dâmbovița (zona Târgoviște, cercetată prin 64 de foraje hidrogeologice; **tabelul 2.13, Fig.2.44**), pentru un model numeric în diferențe finite.

Tabelul 2.13. Forajele de explorare ale acviferului în zona Târgoviște

Foraj	X [m]	Y [m]	T [m ² /zi]	Foraj	X [m]	Y [m]	T [m ² /zi]
1	17215.9	4809.0	1395.0	100	14488.6	10844.8	928.0
2	17386.4	4907.2	1466.0	101	14545.5	10746.7	1027.0
3	17613.6	5005.3	1531.0	102	14715.9	10648.5	634.0
4	17784.1	5103.4	1255.0	104	18977.3	9667.1	320.0
5	17784.0	5103.4	773.0	105	19375.0	9029.2	221.0
6	18693.2	4759.9	814.0	107	26761.4	6281.2	576.0
7	18893.2	4907.2	721.0	108	25909.1	7458.9	277.0
8	19034.1	5005.3	818.0	109	12159.1	5103.4	873.0
9	19034.1	5496.0	285.0	110	12102.3	5005.3	2304.0
10	19602.3	5643.2	1172.0	111	12045.5	4858.1	1067.0
11	19829.5	5790.5	1222.0	112	11647.7	4809.0	941.0
12	20056.8	5888.6	688.0	116	19375.0	13592.8	373.0
13	20340.9	6035.8	1047.0	117	20568.2	13396.6	136.0
14	20568.2	6183.0	18.5	119	23863.6	9618.0	319.0
15	20738.6	6281.2	1677.0	120	23181.8	8293.1	60.0
16	20909.1	6379.3	962.0	124	21818.2	3827.6	271.0
17	21136.4	6477.5	1525.0	128	18920.5	2600.8	12.5
18	21306.8	6673.7	1820.0	131	21875.0	10354.1	101.0
19	21477.3	6968.2	1113.0	132	22670.5	10844.8	85.0
20	21647.7	7017.2	698.0	133	18352.3	294.4	1339.0
21	21590.9	7262.6	1117.0	134	19488.6	883.3	3456.0
22	21875.0	7409.8	1810.0	135	20625.0	2208.2	85.0
23	22159.1	7508.0	470.0	138	25454.5	5790.5	176.0
24	22272.7	7557.0	777.0	154	16193.2	12562.3	580.0
25	22500.0	7753.3	645.0	155	14886.4	8636.6	420.0
26	16875.0	4122.0	835.0	156	8750.0	7655.2	460.0
27	16761.4	3974.8	255.0	157	9034.1	6771.9	398.0
29	16818.2	3631.3	969.0	159	11022.7	6771.9	880.0
83	14318.2	11531.8	760.0	160	10852.3	5692.3	143.0
84	17386.4	10452.3	629.0	161	13465.9	4956.2	233.0
86	17443.2	10943.0	1170.0	162	14715.9	4220.2	198.0
88	14772.7	10648.5	177.0	163	15000.0	3435.0	236.0

Rezolvare:

Modelarea matematică a dinamicii acviferelor are ca obiectiv evaluarea principalelor caracteristici ale curgerii (direcții de curgere, viteze de curgere) pe baza parametrilor hidrogeologici determinați prin foraje hidrogeologice.

Inițializarea distribuției transmisivității pentru modelul în diferențe finite al acviferului freatic din interfluviul Ialomița-Dâmbovița este prima operațiune care se realizează după schematizarea domeniului spațial în care se studiază curgerea.

Domeniul spațial în care este construit modelul (delimitat cu linie grosă continuă în Fig.2.44) este acoperit cu o rețea pătratică de 116 celule cu dimensiunea 1000 x 1000 m.

Inițializarea distribuției transmisivității pentru modelul în diferențe finite presupune calculul valorii medii a transmisivității acviferului în fiecare din cele 116 celule pe baza valorilor din cele 64 de foraje de cercetare (reprezentate prin triunghiuri negre în Fig.2.44).

Calculul celor 116 valori ale transmisivității se realizează prin *kriging zonal*, utilizându-se modelul sferic de variogramă al transmisivității (Fig.2.45), cu parametri:

- Efect de pepită = 167200;
- Palier = 285000;
- Raza de influență = 5850 m

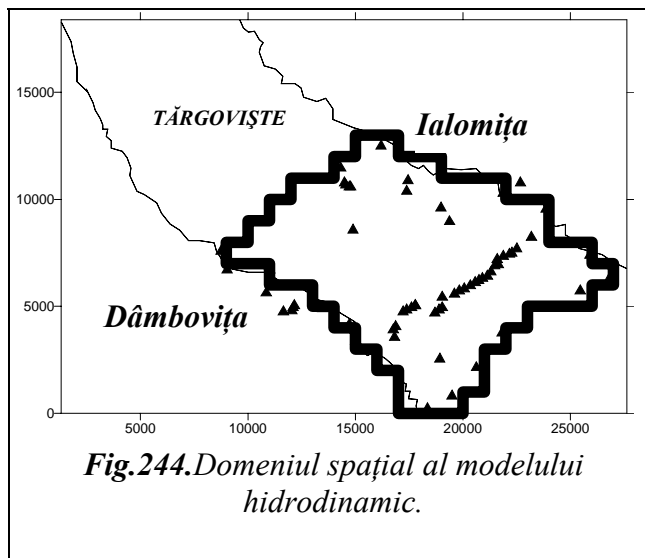


Fig.2.44. Domeniul spațial al modelului hidrodinamic.

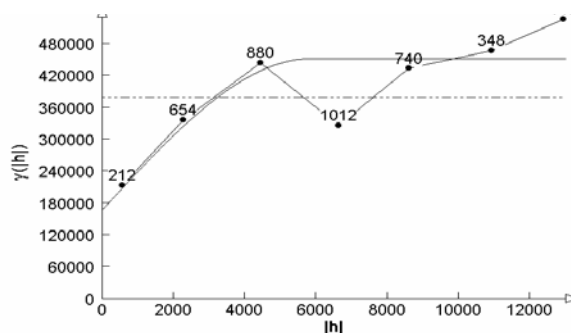


Fig.2.45. Variograma omnidirecțională a transmisivității

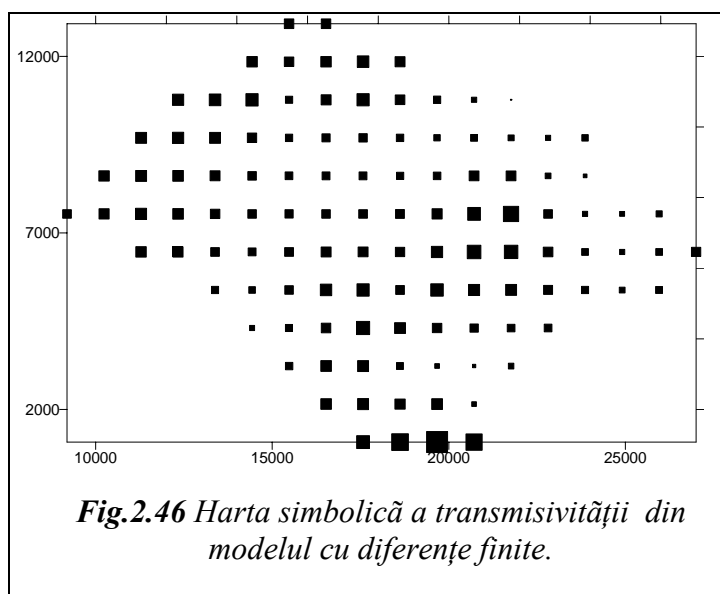


Fig.2.46 Harta simbolică a transmisivității din modelul cu diferențe finite.

Pentru calculul valorii medii a transmisivității, pe fiecare celulă s-au utilizat 16 puncte de discretizare a celulelor rețelei de diferențe finite, pentru a reduce la maximum erorile de estimare zonală.

Distribuția celor 116 valori medii ale transmisivității (tabelul 2.14) se poate reprezenta sub forma unei hărți punctuale cu simboluri având dimensiunile proporționale cu valoarea transmisivității (Fig.2.46) sau hărți simbolice cu tonuri de gri

Tabelul 2.14. Rezultatele inițializării parametrice a modelului numeric

X [m]	Y [m]	T [m ² /zi]	X [m]	Y [m]	T [m ² /zi]	X [m]	Y [m]	T [m ² /zi]
19666.7	1076.9	2876.0	12333.3	6461.5	673.0	22809.5	4307.7	410.6
18619.0	1076.9	1793.6	18619.0	2153.9	671.8	15476.2	8615.4	405.3
20714.3	1076.9	1719.7	18619.0	11846.2	660.2	21761.9	4307.7	403.9
21761.9	7538.5	1481.0	13381.0	8615.4	654.9	19666.7	8615.4	401.7
20714.3	6461.5	1196.8	16523.8	10769.2	654.3	18619.0	9692.3	383.8
21761.9	6461.5	1147.8	16523.8	6461.5	647.2	15476.2	3230.8	380.3
17571.4	4307.7	1127.4	17571.4	6461.5	645.4	19666.7	10769.2	375.1
17571.4	1076.9	1062.2	21761.9	8615.4	641.2	15476.2	9692.3	372.1
20714.3	7538.5	1040.7	20714.3	8615.4	637.3	23857.1	5384.6	366.2
19666.7	5384.6	1002.7	18619.0	10769.2	618.2	18619.0	8615.4	365.9
17571.4	5384.6	994.9	16523.8	4307.7	613.6	15476.2	10769.2	361.4
14428.6	10769.2	983.0	22809.5	6461.5	610.7	13381.0	5384.6	360.6
17571.4	10769.2	961.8	16523.8	12923.1	605.2	15476.2	4307.7	346.9
13381.0	10769.2	884.5	15476.2	12923.1	594.8	25952.4	5384.6	346.2
17571.4	11846.2	870.8	13381.0	7538.5	591.1	18619.0	3230.8	343.3
16523.8	5384.6	865.5	14428.6	9692.3	580.8	23857.1	6461.5	329.6
12333.3	10769.2	833.0	15476.2	11846.2	578.4	20714.3	9692.3	328.9
11285.7	7538.5	818.4	19666.7	4307.7	577.1	25952.4	6461.5	328.8
19666.7	6461.5	814.9	18619.0	6461.5	576.6	23857.1	9692.3	308.7
20714.3	5384.6	802.5	27000.0	6461.5	560.6	14428.6	5384.6	306.7
12333.3	9692.3	801.0	22809.5	5384.6	550.2	19666.7	9692.3	294.4
17571.4	2153.9	792.8	18619.0	7538.5	521.8	25952.4	7538.5	282.4
17571.4	3230.8	782.0	18619.0	5384.6	509.7	21761.9	9692.3	268.8
11285.7	9692.3	779.4	15476.2	6461.5	507.8	24904.8	5384.6	258.9
11285.7	8615.4	778.8	13381.0	6461.5	506.0	22809.5	8615.4	253.7
21761.9	5384.6	773.6	17571.4	7538.5	502.2	21761.9	3230.8	246.2
13381.0	9692.3	767.8	9190.5	7538.5	499.3	20714.3	10769.2	225.4
19666.7	2153.9	757.2	16523.8	7538.5	495.2	24904.8	6461.5	217.3
12333.3	8615.4	756.3	14428.6	8615.4	494.7	23857.1	7538.5	215.6
18619.0	4307.7	755.2	22809.5	7538.5	494.5	22809.5	9692.3	211.5
16523.8	2153.9	738.4	15476.2	5384.6	485.2	14428.6	4307.7	208.8
16523.8	3230.8	737.8	14428.6	7538.5	482.8	24904.8	7538.5	205.8
12333.3	7538.5	729.6	17571.4	9692.3	476.0	20714.3	2153.9	192.9
11285.7	6461.5	718.6	20714.3	4307.7	462.3	19666.7	3230.8	191.0
16523.8	11846.2	714.7	15476.2	7538.5	461.0	23857.1	8615.4	150.7
14428.6	11846.2	713.3	16523.8	9692.3	445.6	20714.3	3230.8	136.8
10238.1	8615.4	702.8	14428.6	6461.5	433.1	21761.9	10769.2	101.8
19666.7	7538.5	690.6	16523.8	8615.4	420.1			
10238.1	7538.5	679.0	17571.4	8615.4	419.1			

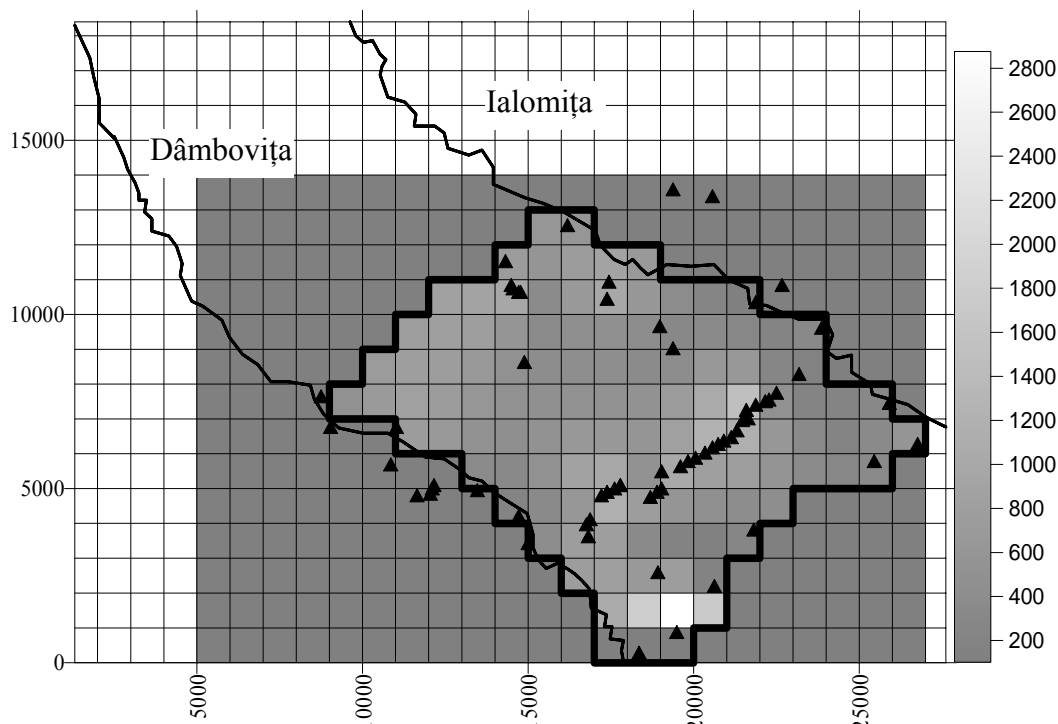


Fig.2.47.Harta simbolică a transmisivității acviferului freatic în modelul cu diferențe finite

sau culori (**Fig.2.47**). Pe harta simbolică sunt suprapuse: rețeaua de discretizare și cele 64 foraje hidrogeologice de explorare.

Programele de calcul utilizate pentru modelarea matematică a proceselor dinamice și chimice din sisteme acvifere (ex.: Visual Modflow, SUTRA etc.) au posibilitatea de a prelua valorile transmisivităților calculate cu programe de tip SURFER și salvate în fișierele de tip rețea (GRD).

COMENTARIU

Modelarea matematică a proceselor geologice se bazează pe legități fizico-chimice fundamentale, pe un domeniu spațial precis localizat și o distribuție parametrică specifică.

Cea mai performantă metodă actuală de estimare a distribuției parametrice specifice pentru domeniul spațial al unui model numeric este kriging-ul zonal!

Evaluarea distribuției transmisivității pentru acviferul cu nivel liber din interfluviul Ialomița-Dâmbovița este un exemplu de aplicare corectă a metodologiei de inițializare parametrică a unui model în diferențe finite cu valorile unuia (transmisivitatea acviferului) din multitudinea de parametri necesari acestuia. Metoda de interpolare utilizată, kriging-ul zonal, este aceeași utilizată pentru calculul rezervelor și este cea mai performantă metodă de interpolare.

Analog, se pot inițializa pentru orice model matematic parametri precum: temperatura, concentrația diferitelor substanțe poluante în aer sau sol, starea de eforturi în masivele de roci, conținuturile în metale ale filoanelor etc.

Avantajul utilizării kriging-ului ca metodă de interpolare constă în posibilitatea evaluării erorilor de estimare asociate distribuției estimate a parametrilor.

3. ERORILE ESTIMĂRII DISTRIBUȚIEI SPAȚIALE

Estimarea distribuției spațiale a unei variabile este afectată în mod inevitabil de erori din cauza informației incomplete. Extraordinara variabilitate spațială a caracteristicilor unui *proces natural* nu poate fi descrisă fără erori decât dacă se fac *măsurători în toate punctele domeniului spațial* în care acesta este studiat.

Dacă am ști prin *măsurători* ce caracteristici are procesul studiat în fiecare punct al domeniului studiat *obiectul estimării geostatistice ar dispărea și geostatisticianul de asemenea*. Din fericire pentru geostatisticieni acest lucru nu este posibil, din cauza costurilor prohibitive ale unei probări exhaustive.

Disponând de un *număr limitat de valori* ale caracteristicii studiate trebuie să adoptăm un model de variabilitate spațială (*variograma*) despre care am văzut că este aproximativ, să acceptăm ipoteze de lucru care sunt adaptate mai mult sau mai puțin datelor prelucrate (*normalitatea, staționaritatea etc.*), să ne asumăm un risc în estimările realizate (α - *riscul erorii de genul I*).

Studiul erorilor de estimare a distribuției spațiale are ca obiectiv reducerea acestora la minimum și se realizează în două etape distincte:

- *calculul erorilor*, etapă în care se utilizează kriging-ul pentru a extrage maximum de informație din datele disponibile;
- *reducerea erorilor calculate* în etapa anterioară prin completarea datelor disponibile cu altele suplimentare.

3.1. CALCULUL ERORILOR

Este obligatoriu ca studiul erorilor să debuteze cu calculul acestora. Fazele acestei etape de prelucrare sunt următoarele:

- identificarea factorilor care determină mărimea erorii de estimare;
- definirea metodelor de reducere a erorilor de estimare;
- cuantificarea erorilor de estimare

Factorii care determină mărimea erorii de estimare.

Principali factori care influențează mărimea erorii de estimare sunt: *numărul valorilor* utilizate pentru estimare, *dispoziția spațială a punctelor de observație* din vecinătatea ariei de estimare, *natura variabilității* (modelul de variogramă) caracteristicii studiate și *suportul estimării*.

La nivel calitativ este evident că eroarea de estimare într-un punct p (**Fig.3.1**) va fi maximă în cazul utilizării unei singure valori (**Fig.3.1a**), redusă în cazul utilizării a patru valori grupate, spre exemplu, în NV-ul punctului p (**Fig.3.1b**) și minimă în cazul plasării punctului p în centrul unui cerc pe circumferința căruia sunt plasate cele patru valori (**Fig.3.1c**).

Natura variabilității caracteristicii studiate (cuantificată în *modelul de variogramă*) este un factor de mare complexitate. Dacă se operează cu variabile cu o bună continuitate și variabilitate redusă estimările sunt afectate de erori mici. Trebuie însă avut în vedere că natura variabilității se poate

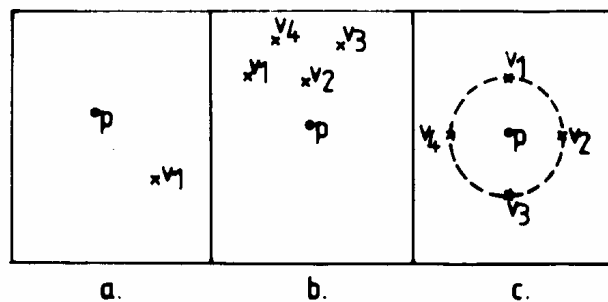


Fig. 3.1. Efectul numărului și distribuției punctelor asupra erorii de estimare

modifica de la o zonă la alta a suprafeței cercetate.

În practică, deseori se utilizează același model de variogramă pentru întreaga suprafață cercetată, ceea ce în cazul neomogenității structurale este neadecvat. Chiar dacă modelul de variabilitate se păstrează, multe categorii de parametri prezintă *efectul de proporționalitate* (corelația între varianță și valoarea parametrului) care afectează și el semnificativ eroarea de estimare.

Asociați și cu efectul *suportului de estimare* care implică o tratare detaliată, toți acești factori interacționează într-un mod complex care trebuie analizat pentru fiecare estimare în parte, situațiile fiind diversificate:

- pentru parametri cu o bună continuitate și variabilitate redusă este preferabil unui număr mare de valori plasate la mare distanță de punctul de estimare, un număr redus de valori măsurate în imediata vecinătate a acestuia;
- pentru parametri cu continuitate redusă și variabilitate accentuată este preferabil să se dispună de mai multe valori măsurate, plasate la distanță mare de punctul de estimare;
- pentru parametri cu o bună continuitate și variabilitate redusă valorile măsurate pe locații apropiate sunt redundante în timp ce pentru parametri cu variabilitate accentuată aceleași valori pot să nu fie redundante.

Toate aceste considerații de natură calitativă trebuie avute în vedere la utilizarea diferitelor instrumente de cuantificare a *erorilor de estimare*.

Influența dimensiunii *suportului estimării* este proporțională cu diferența dintre volumul spațial asociat *valorilor măsurate* și volumul spațial asociat *valorii estimate*. Eliminarea totală a efectului dimensiunii suportului valorilor măsurate asupra estimărilor presupune egalitatea între dimensiunea suportului acestora și suportul estimărilor.

În cazul în care valorile măsurate au *suport punctual* (ex.: grosimea unui strat de argilă măsurată într-un foraj, conținutul de brom determinat într-o probă de apă recoltată dintr-un puț, sarcina piezometrică măsurată în forajele hidrogeologice etc.) iar *estimările sunt punctuale*, influența suportului este nulă.

Efectul suportului intervine semnificativ când *pe baza unor valori punctuale* (ex.: valori ale proprietăților fizico-chimice determinate în probe prelevate din foraje) *se estimează valoarea medie pe o suprafață* (ex.: delimitarea zonei cu anumite caracteristici chimice maxime admise de STAS-ul de potabilitate a apei).

Situații similare apar atunci când pe baza parametrilor hidrogeologici (transmisivitate, difuzivitate hidraulică etc.) determinați prin testări hidro dinamice se inițializează parametric modelele de simulare numerică a dinamicii acviferelor. În aceste situații suportul asociat valorilor parametrilor determinați este aria de influență a pompărilor executate, diferită de suprafața elementelor utilizate pentru discretizarea acviferului modelat.

Eliminarea erorilor introduse de diferența dintre suportul valorilor măsurate și cel al estimărilor presupune adoptarea unei funcții care să exprime efectul dimensiunii suportului asupra distribuției valorilor estimate. Alegerea acestei funcții se face ținând seama de particularitățile structurii variabilei studiate.

Dacă de cele mai multe ori valorile măsurate sunt asociate unor *suporturi punctuale*, în estimare, pentru creșterea reprezentativității evaluărilor și reducerea volumului de calcule, se recurge la *suporturi areale*. Creșterea dimensiunii suportului valorilor utilizate în estimările geostatistice are ca rezultat:

- menținerea valorii medii a selecției de date primare;

- reducerea amplitudinii selecției de valori utilizate pentru estimare prin reducerea valorii maxime;
- creșterea valorii minime;
- reducerea dispersiei;
- reducerea asimetriei (diferența dintre medie și mediană).

Reducerea *dispersiei* și *asimetriei* depinde de gradul de continuitate și particularitățile structurale ale variabilelor:

- pentru variabile cu structurare spațială slabă (variogramă efect de pepită total din **Fig.3.2a**) creșterea suportului conduce rapid la reducerea asimetriei (**Fig.3.2b, c**);
- pentru variabile cu o structurare spațială bună (**Fig.3.3a**) efectul creșterii suportului asupra dispersiei și asimetriei este redus (**Fig.3.3b, c**).

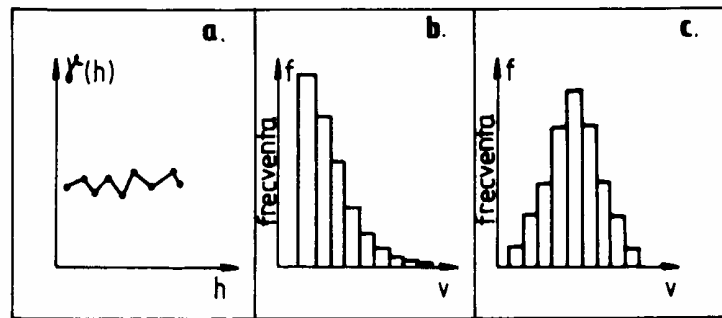


Fig. 3.2. Efectul creșterii suportului asupra distribuției valorilor necorelate spațial:

a - variograma valorilor punctuale;

b - histograma valorilor punctuale;

c - histograma valorilor medii pe blocuri de 10x10

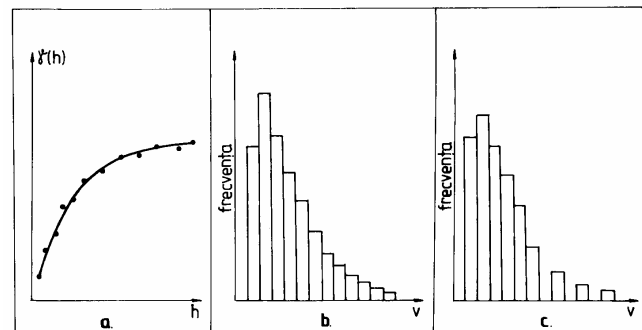


Fig. 3.3. Efectul creșterii suportului asupra distribuției valorilor corelate spațial:

a - variograma valorilor punctuale;

b - histograma valorilor punctuale;

c - histograma valorilor medii pe blocuri (10 x 10)

Variograma este un instrument de mare eficiență în evaluarea efectului dimensiunii suportului, singura informație care *scapă* variogramei fiind cea a modului de corelare a valorilor extreme, cu influență semnificativă însă asupra modificării gradului de simetrie al distribuției.

Modalitățile de reducere a erorilor de estimare.

Există două categorii de metode de reducere a erorilor de estimare cu aplicabilitate generală: cele de *corecție a efectului de suport* și de *corecție a distribuției frecvențelor valorilor* prelucrate.

Metodele de corecție a efectului de suport au două caracteristici principale: păstrează neschimbată valoarea medie de selecție și ajustează varianța selecției cu un anumit factor care reprezintă raportul dintre varianța proprie suportului de estimare și varianța totală a selecției. Ele diferă prin modul în care acționează asupra simetriei selecțiilor și sunt alese în funcție de particularitățile spațiale ale variabilei eșantionate, parțial conținute în modelul de variogramă:

- dacă valorile au o slabă corelație spațială se alege un procedeu de corecție care reduce asimetria proporțional cu creșterea dimensiunii suportului;
- dacă valorile sunt bine corelate se aleg proceduri de corecție care nu modifică simetria selecției de valori.

Corecția afină este unul din cele mai simple procedee de corectare a efectului de suport care *nu modifică simetria*. Ideea de bază este că varianța distribuției trebuie redusă prin gruparea valorilor selecției în jurul mediei. Corecția afină transformă o valoare a unei distribuții v în alta, v' , utilizând formula lineară:

$$v' = \sqrt{f} \cdot (v - m) + m, \quad (3.1)$$

în care:

m - media ambelor selecții (neschimbată în urma corecției afine);

f - factorul de transformare al varianței;

Dacă σ^2 este varianța primei distribuții, varianța valorilor transformate utilizând factorul f este:

$$(\sigma')^2 = f \cdot \sigma^2 \quad (3.2)$$

Utilizând transformarea (3.1) este evident că diagrama $v-v'$ este o dreaptă, deci prin corecția afină nu se modifică nici legea de distribuție a valorilor și nici simetria ei. Corecția afină este recomandat să fie utilizată cu factori de corecție mari și pentru *valori de prag* apropiate de valoarea medie.

Corecția lognormală modifică simetria distribuției inițiale și este aplicabilă numai pentru valori cu distribuție lognormală. Ideea de bază este că efectul creșterii suportului este similar cu reducerea varianței într-o distribuție lognormală. Acest efect se materializează în reducerea gradului de asimetrie, în condițiile menținerii mediei de selecție. Formula de transformare în cazul corecției logaritmice este:

$$v' = a \cdot v^b, \quad (3.3)$$

în care coeficienții a și b sunt calculați cu relațiile:

$$a = \frac{m}{\sqrt{f \cdot c_v^2 + 1}} \cdot \left[\frac{\sqrt{c_v^2 + 1}}{m} \right]^b \quad (3.4)$$

și

$$b = \sqrt{\frac{\ln(f \cdot c_v^2 + 1)}{\ln(c_v^2 + 1)}}, \quad (3.5)$$

în care c_v este coeficientul de variație.

Deoarece coeficientul a calculat cu relația (3.4) păstrează media neschimbată numai în cazul unei distribuții lognormale a valorilor originale, orice abatere de la distribuția lognormală trebuie corectată cu relația:

$$v'' = \frac{m}{m'} \cdot v', \quad (3.6)$$

în care m' este media valorilor transformate. Dacă distribuția valorilor originale este lognormală, raportul celor două medii este unitar. Prin corecția lognormală crește simetria valorilor transformate o dată cu reducerea varianței datorată creșterii suportului. Corecția este sensibilă la valori de prag extreme.

Factorul de corecție este raportul dintre varianța valorilor variabilei măsurate în interiorul suportului de estimare (B) și varianța totală calculată pe toată suprafața (A) în funcție de suportul valorilor originale (P):

$$f = \frac{\sigma^2(B, A)}{\sigma^2(P, A)}, \quad (3.7)$$

în care:

B - suportul de estimare;

P - suportul valorilor măsurate (în general punctual);

A - suprafața totală explorată (**Fig.3.4**).

Dispersia selecției de valori ($\sigma^2(P, A)$) are două componente în cazul separării suprafeței în mai multe blocuri (B_1, B_2, B_3, B_4 în **Fig.3.4**):

- varianța în interiorul blocurilor

($\sigma^2(P, B)$);

- varianța dintre blocuri ($\sigma^2(B, A)$),

între care există relația:

$$\sigma^2(P, A) = \sigma^2(P, B) + \sigma^2(B, A). \quad (3.8)$$

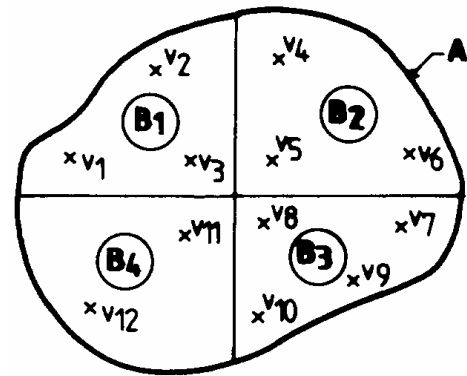


Fig. 3.4. Descompunerea varianței totale

În mod clasic, componentele varianței pentru exemplul din **Fig.3.4** se calculează după relațiile cunoscute:

$$\sigma^2(P, A) = \frac{1}{12} \sum_{i=1}^{12} (v_i - m)^2, \quad \text{în care } m = \frac{1}{12} \sum_{i=1}^{12} v_i \quad (3.9)$$

$$\sigma^2(P, B) = \frac{1}{12} \left\{ [(v_1 - m_{B1})^2 + (v_2 - m_{B1})^2 + (v_3 - m_{B1})^2] + [(v_4 - m_{B2})^2 + (v_5 - m_{B2})^2 + (v_6 - m_{B2})^2] + [(v_7 - m_{B3})^2 + (v_8 - m_{B3})^2 + (v_9 - m_{B3})^2 + (v_{10} - m_{B3})^2] + [(v_{11} - m_{B4})^2 + (v_{12} - m_{B4})^2] \right\} \quad (3.10)$$

și

$$\sigma^2(B, A) = \frac{1}{4} \left[(m_{B1} - m)^2 + (m_{B2} - m)^2 + (m_{B3} - m)^2 + (m_{B4} - m)^2 \right], \quad (3.11)$$

în care:

$$m_{B1} = \frac{1}{3} (v_1 + v_2 + v_3) \quad (3.12)$$

$$m_{B2} = \frac{1}{3}(v_4 + v_5 + v_6) \quad (3.13)$$

$$m_{B3} = \frac{1}{3}(v_7 + v_8 + v_9 + v_{10}) \quad (3.14)$$

$$m_{B4} = \frac{1}{3}(v_{11} + v_{12}) \quad (3.15)$$

Varianța pentru orice suprafață, adică pentru orice suport, poate fi estimată din variograma de a cărei relație de definiție este direct legată:

$$\sigma^2(P, A) = \frac{1}{n} \sum_{i=1}^n (v_i - m)^2 = \frac{1}{n} \sum_{i=1}^n v_i^2 - m^2 \quad \text{cu} \quad m = \frac{1}{n} \sum_{i=1}^n v_i \quad (3.16)$$

Rezultă că :

$$\begin{aligned} \sigma^2(P, A) &= \frac{1}{n} \sum_{i=1}^n v_i^2 - \left(\frac{1}{n} \sum_{i=1}^n v_i \right)^2 = \frac{1}{2n} \sum_{i=1}^n v_i^2 + \frac{1}{2n} \sum_{j=1}^n v_j^2 - \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n v_i v_j = \\ &= \frac{1}{2n^2} \sum_{i=1}^n \sum_{j=1}^n v_i^2 + \frac{1}{2n^2} \sum_{j=1}^n \sum_{i=1}^n v_j^2 - \frac{2}{2n^2} \sum_{i=1}^n \sum_{j=1}^n v_i v_j = \\ &= \frac{1}{2n^2} \left(\sum_{i=1}^n \sum_{j=1}^n v_i^2 + \sum_{j=1}^n \sum_{i=1}^n v_j^2 - 2 \sum_{i=1}^n \sum_{j=1}^n v_i v_j \right) = \frac{1}{2n^2} \sum_{i=1}^n \sum_{j=1}^n (v_i - v_j)^2 \end{aligned} \quad (3.17)$$

care este similară cu relația de definiție a variogramei :

$$\gamma(h) = \frac{1}{2N(h)} \sum_{(i,j)_{h_{ij}=h}} (v_i - v_j)^2 \quad (3.18)$$

Varianța și variograma sunt medii ale unor diferențe de pătrate pentru toate perechile de valori separate de o anumită distanță.

Varianța poate fi considerată ca un fel de variogramă care ia în calcul perechile de valori pentru care h_{ij} este conținut în suprafața explorată (A din **Fig.3.4**):

$$\sigma^2(P, A) = \frac{1}{2N(A)} \sum_{(i,j)_{h_{ij}=h}} (v_i - v_j)^2 \quad (3.19)$$

Pentru calculul varianței pe baza variogramei se parcurg următoarele etape de prelucrare:

- Calculul variogramei experimentale pe baza tuturor valorilor primare disponibile.
- Modelarea variogramei experimentale.
- Discretizarea suprafeței fiecărui bloc în care este împărțită suprafața totală. Numărul de puncte de discretizare (n) este în funcție de extinderea blocului și de numărul valorilor măsurate.
- Medierea celor n^2 valori ale variogramei modelate cu relația:

$$(\sigma^*(P,A))^2 = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \gamma^*(h_{ij}); h_{ij} \in A \quad (3.20)$$

Cu ecuația (3.20) pot fi calculate cele două componente ale varianței necesare estimării factorului de corecție definit de relația (3.7):

- $\sigma^2(P,B)$ se calculează prin discretizarea fiecărui bloc prin câteva puncte și calculul mediei valorilor variogramelor obținute din model;
- $\sigma^2(P,A)$ poate fi calculat cu relația clasică (8.9) sau cu ajutorul variogramei.

Metodele de corecție a efectului distribuției frecvenței valorilor reduc erorile asociate estimărilor locale sau zonale realizate prin kriging. Erorile de estimare sunt condiționate de cunoașterea *distribuției valorilor estimate* (necunoscute).

Instrumentele utilizate pentru determinarea distribuției valorilor estimate sunt aceleași utilizate pentru estimările punctuale sau zonale, dar aplicate valorilor măsurate, transformate în *indicatori*.

Există două categorii de metode pentru estimarea funcției de distribuție a valorilor estimate:

- *metode neparametrice*, cu o descriere incompletă a distribuției prin calculul *funcției de frecvență cumulată* ($F(v)$) pentru diferite valori ale variabilei v ;
- *metode parametrice* care utilizează *modelul funcției aleatoare* și permit determinarea unei funcții de distribuție pentru orice valoare a variabilei v .

Cea mai grosieră metodă pentru estimarea distribuției valorilor necunoscute (estimate) este utilizarea *histogramei valorilor măsurate*. Această variantă ignoră complet posibilitatea modificării caracterului distribuției prin contribuția valorilor necunoscute și influența geometriei punctelor de probare disponibile.

Eliminarea influenței grupării neadecvate a punctelor de probare poate fi realizată prin estimarea punctuală a valorilor variabilei studiate în nodurile unei rețele regulate. Această metodă conduce la reducerea variabilității reale a variabilei studiate prin *netezirea* proprie metodelor de interpolare geostatistice. Gradul de netezire al estimărilor este proporțional cu numărul de valori utilizate în combinațiile lineare.

Soluția problemei estimării corecte a proporției pe care o reprezintă valorile aflate sub o anumită valoare de prag se află în înțelegerea modului în care se procedează când avem la dispoziție o selecție exhaustivă asupra variabilei v .

Pentru valoarea v_{p_1} , valoarea funcției de distribuție se calculează cu relația:

$$F(v_{p_1}) = \frac{Nv_{p_1}}{n}, \quad (3.21)$$

în care:

- Nv_{p_1} - numărul de valori mai mici decât valoarea v_{p_1} ;
- n - numărul total de valori ale variabilei v .

Introducem noțiunea de *variabilă indicatoare* (i) sub forma:

$$i_j = \begin{cases} 1, & \text{dacă } v_j \leq v_{p_1} \\ 0, & \text{dacă } v_j > v_{p_1} \end{cases} \quad (3.22)$$

Calculul erorilor

Utilizând această convenție harta punctuală a variabilei se transformă într-o hartă indicatoare cu două simboluri: 0 și 1. Numărul valorilor mai mici ca valoarea de prag v_{p1} se calculează cu relația:

$$N_{v_{p1}} = \sum_{j=1}^n i_j, \quad (3.23)$$

iar proporția acestora în numărul total de valori cu relația:

$$F(v_{p1}) = \frac{\sum_{j=1}^n i_j}{n}. \quad (3.24)$$

Acest procedeu se repetă pentru orice v_p transformându-se valorile v_1, v_2, \dots, v_n în indici: $i_1(v_p), \dots, i_n(v_p)$ cu relația generală:

$$i_j(v_p) = \begin{cases} 1, & \text{dacă } v_j \leq v_p \\ 0, & \text{dacă } v_j > v_p \end{cases} \quad (3.25)$$

iar funcția de frecvență cumulată se calculează cu relația:

$$F(v_p) = \frac{1}{n} \sum_{j=1}^n i_j(v_p). \quad (3.26)$$

Funcția de frecvență cumulată (**Fig.3.5**), pentru valori mai mici decât valoarea de prag v_{p1} , este zero, la v_{p2} sare la $1/n$ și continuă cu salturi de $1/n$ până la valoarea de prag v_{pk} când ia valoarea 1 dacă v_{pk} este mai mare sau egală cu valoarea maximă a selecției.

În acest mod se transferă noțiunea de proporție a valorilor aflate sub un anumit prag în cea de medie a unui *indicator* care permite adaptarea metodelor de estimare pentru precizarea completă a distribuției valorilor necunoscute.

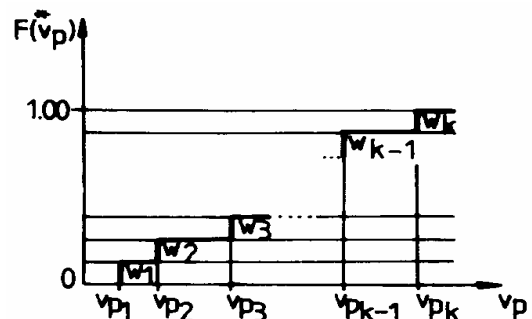


Fig. 3.5. Funcție de frecvență pentru rețea de probare uniformă

Estimarea distribuției cumulative globale este necesară pentru evaluarea preciziei estimărilor pe suprafețe mari. Pentru o *valoare de prag* v_p orice variabilă continuă V poate fi transformată într-un *indicator* $I(v_p)$ utilizând ecuația (3.25).

Deoarece în practică nu avem acces la o selecție exhaustivă, se poate realiza o estimare a proporției reale a valorilor plasate sub o anumită valoare de prag printr-o mediere ponderată a indicatorului:

$$F^*(v_p) = \sum_{j=1}^n w_j i_j(v_p), \text{ cu } \sum_{j=1}^n w_j = 1 \quad (3.27)$$

Estimatorul funcției frecvenței cumulative globale este format din n termeni cu ponderi diferite în funcție de distribuția spațială a punctelor de observație. (în raport cu valorile variabilei cercetate.) Dacă valorile disponibile $v_1, v_2, v_3, \dots, v_n$ sunt plasate într-o rețea regulată ponderile vor fi egale $w_1 = w_2 = \dots = w_n = 1/n$.

Dacă rețeaua de observație este *neregulată* și punctele sunt localizate preferențial în zone cu *valori mari*, ponderea valorilor mici va fi *subestimată*. În caz contrar, dacă punctele sunt plasate în zone cu *valori mici*, ponderile vor fi *supraestimate*.

Când se estimează proporția globală a valorilor plasate sub o anumită valoare de prag dintr-o selecție de date *clusterizate* este dificil de extras un set cu distribuție regulată și se preferă ponderarea valorilor disponibile prin *metoda declustering-ului poligonal sau celular*. În această variantă saltul funcției cumulative la fiecare valoare de prag este diferit și diferit de $1/n$ (**Fig.3.6**).

Pentru descrierea distribuției cumulative globale se utilizează în mod suplimentar varianța de dispersie $(\sigma^*)^2$ și coeficientul de asimetrie (β^*) calculate cu relațiile clasice:

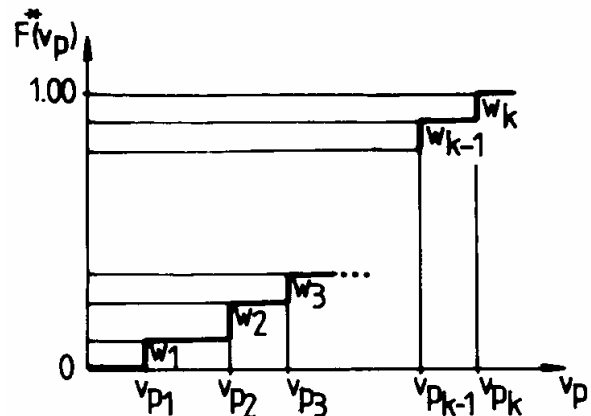


Fig. 3.6. Funcție de frecvență pentru rețea de probare neuniformă

$$(\sigma^*)^2 = \sum_{j=1}^n w_j (v_j - m^*)^2 \quad (3.28)$$

și:

$$\beta = \frac{\sum_{j=1}^n w_j (v_j - m^*)^3}{(\sigma^*)^3} \quad (3.29)$$

în care:

m^* - media estimată a valorilor măsurate;

w_j - ponderile estimate prin declustering poligonal/celular.

Estimarea distribuției cumulative locale este necesară pentru evaluarea preciziei estimărilor zonale. Cunoașterea distribuției locale este deosebit de importantă în studiile ambientale pentru estimarea corectă a concentrației de poluant în ariile de interes. Proporția valorilor situate sub o anumită valoare de prag fiind media indicilor, este posibilă utilizarea indicatorilor valorilor disponibile din vecinătatea ariei de interes pentru estimarea funcției distribuției locale.

În **Fig.3.7** este desenată o suprafață de formă neregulată explorată în interior și exterior prin zece puncte de observație în care sunt măsurate valorile unei variabile. Pentru estimarea frecvenței corespunzătoare a două valori de prag, $v_{p1} = 190$ și

$v_{p2} = 200$, este exemplificat modul de calcul al funcțiilor de frecvență $F^*(190) = 0,50$ și $F^*(200) = 0,40$.

Alegerea

valorilor de prag este determinată nu numai de realizarea curbei frecvențelor cumulate (care nu este un scop în sine), ci și de obiectivele estimării. În evaluarea calității apei unui acvifer, valorile de prag pot fi limitele maxime și minime ale anumitor parametri fizico-chimici ai apei. Pentru evaluarea rezervelor unui zăcământ, valoarea de prag

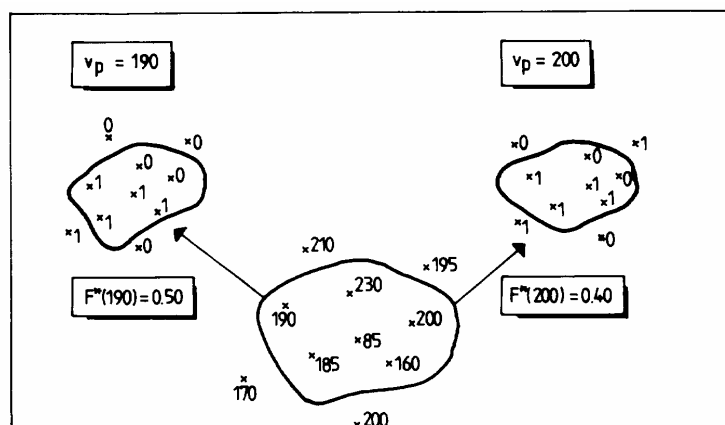


Fig. 3.7. Estimarea frecvenței cumulate locale pentru două valori de prag

poate fi valoarea conținutului minim exploatabil. Dacă nu există astfel de rațiuni practice care să determine alegerea valorilor de prag, acestea se aleg în număr de 9 corespunzător celor 9 decile ale selecției studiate.

Pentru estimarea unei curbe complete este necesară interpolarea între punctele estimate și extrapolarea înaintea primului și după ultimul punct estimat. Aceste interpolări și extrapolări presupun adoptarea unor ipoteze în legătură cu distribuția valorilor necunoscute încadrate între două limite clare: zero (valoarea minimă) și unu (valoarea maximă). Realismul interpolărilor și extrapolărilor este asigurat de identificarea corectă, de obicei prin intermediul variogramei indicatoare, a legității de distribuție spațială a variabilei studiate.

Transformarea valorilor variabilelor exprimate în indici, corespunzători diferitelor praguri valorice (v_p), permite identificarea legilor de distribuție spațială a diferitelor categorii de valori.

Utilizarea variogramelor indicatoare stabilite pentru indicii diferitelor praguri valorice este o caracteristică proprie kriging-ului. Metoda poligonală sau metoda inversului distanței atribuie valorilor măsurate aceeași pondere în procesul de interpolare pentru orice prag valoric (v_p).

În kriging-ul ordinar ponderile acordate indicilor din vecinătatea punctelor de estimare sunt dependente de modelul de variogramă corespunzător pragului valoric utilizat.

Sunt multe situații în care modelele de continuitate (variogramele) diferă semnificativ de la o valoare de prag la alta. În zăcămintele de petrol fisurate permeabilitatea ridicată poate corespunde fisurilor, în timp ce permeabilitatea redusă poate fi datorată lentilelor de argilă (Fig. 3.8a). O hartă indicatoare pentru o valoare de prag ridicată a permeabilității poate separa sistemele de fisuri de restul zăcământului, indicând zonele fisurate (permeabilitate maximă) prin zero-uri iar restul zăcământului prin 1 (Fig. 3.8b). Harta indicatoare pentru valori de prag reduse ale permeabilității separă intercalațiile de argile sub forma unor zone marcate prin simbolul 1 (Fig. 3.8c). Tipurile de continuitate din hărțile indicatoare vor reflecta stilul structural al sistemelor de fisuri (Fig. 3.8b) și caracterul depozițional al intercalațiilor argiloase (Fig. 3.8c).

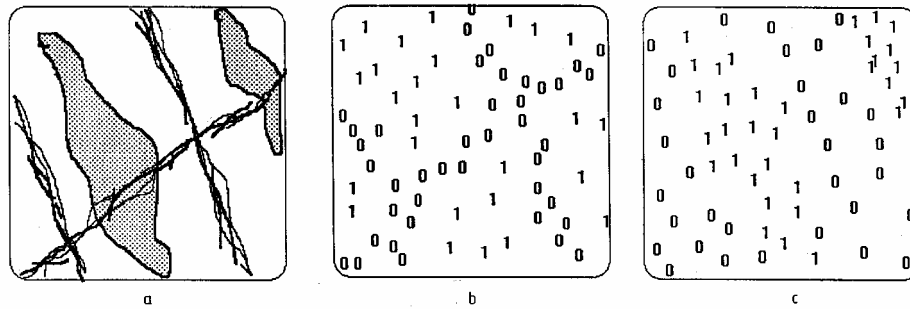


Fig. 3.8. Hărți indicatoare pentru diferite valori de prag: a - zonele fisurate și lentilele de argilă; b - hartă simbolică pentru permeabilitate mare; c - hartă simbolică pentru permeabilitate scăzută.

Variogramele indicatoare experimentale calculate pentru hărțile indicatoare sunt în general mai bine corelate decât variogramele experimentale calculate pe baza valorilor brute măsurabile. Atât timp cât indicii nu pot avea decât valoarea 1 și 0, variograma indicatoare este mai puțin afectată de valorile extreme, posibil aberante, ale variabilei studiate.

Structurile reflectate de variogramele indicatoare pot fi afectate de gruparea spațială preferențială a punctelor de observație. Modelul de variogramă valabil pentru toate valorile de prag este obținut în mod curent din harta indicatoare a unei valori de prag apropiată de valoarea medianei.

Kriging-ul aplicat indicilor și care folosește variograme distincte pentru fiecare valoare de prag poartă denumirea de *kriging indicator*. Numai în cazul în care variogramele tuturor valorilor de prag sunt similare este acceptabilă utilizarea unui singur model de continuitate (variogramă) pentru toate valorile de prag, și anume cel corespunzător valorii de prag egală cu mediana selecției.

Estimările realizate prin *kriging indicator* trebuie să respecte două condiții principale:

- proporția valorilor negative și supraunitare pentru orice valoare de prag trebuie să fie zero. Utilizându-se diferite modele de variogramă funcție de valoarea de prag, estimările kriging-ului indicator nu respectă întotdeauna această condiție, corecția realizându-se prin egalarea cu zero a valorilor negative și cu unu a valorilor supraunitare.

- proporția valorilor estimate sub o anumită valoare de prag v_{p1} trebuie să fie mai mică decât a celor plasate sub o valoare de prag mai mare v_{p2} .

$$F^*(v_{p1}) < F^*(v_{p2}); v_{p1} < v_{p2} \quad (3.30)$$

O modalitate de satisfacere a acestei condiții este utilizarea doar a ponderilor pozitive a căror sumă este unitară și a aceleiași ponderi pentru toate valorile de prag (caracteristic metodei inversului distanței sau metodei poligonale).

În cazul kriging-ului indicator, chiar dacă se utilizează același model de variogramă (variograma indicatoare pentru valoarea de prag egală cu mediana), această a doua condiție nu este îndeplinită. Abaterile introduse fiind reduse ele pot fi eliminate ușor prin procedee matematice simple.

Cuantificarea erorilor de estimare

Toate metodele de cuantificare a erorilor operează cu noțiunea de eroare de estimare definită de relația:

$$e = v^* - v, \quad (3.31)$$

în care:

- v^* - valoarea estimată;
- v - valoarea reală.

Astfel definită, eroarea nu poate fi calculată pentru că *nu se cunoaște valoarea reală în punctul de estimare* (v).

Exprimarea erorii se face prin *indici de incertitudine și intervale de încredere* pentru valorile estimate.

Indicii de incertitudine se stabilesc funcție de factorul care influențează precizia estimării. Valoarea indicilor incertitudinii nu are importanță, indicii fiind utilizați doar pentru compararea erorilor comise la estimarea variabilei în diferite puncte.

Cel mai simplu indice de incertitudine este definit în raport cu numărul de valori utilizate la estimare:

$$I_1 = n_{p_0} \quad (3.32)$$

în care n_{p_0} este numărul de valori măsurate din vecinătatea punctului de estimare (p_0). Valorile estimate (v^*) asociate cu un indice I_1 mare au erori de estimare mici și invers.

Ținându-se seama de distanța dintre punctele de observație și punctul de estimare se definește indicele de incertitudine:

$$I_2 = \frac{1}{\bar{d}}, \quad (3.33)$$

în care \bar{d} este media distanțelor de la punctul de estimare (p_0) la punctele de observație din vecinătate. Pe măsură ce această medie crește, deci indicele I_2 descrește, eroarea de estimare este mai mare.

Deși I_1 și I_2 cuantifică influența a doi factori esențiali ai preciziei de estimare, ei nu pot cuantifica interacțiunea acestor factori, și nici influența structurii spațiale a variabilei asupra rezultatelor probării.

Un al treilea indice al incertitudinii de estimare, care ține seama și de interacțiunea principalilor factori, este:

$$I_3 = \tilde{\sigma}_R^2 = \tilde{c}_{00} + \sum_{i=1}^n \sum_{j=1}^n w_i w_j \tilde{c}_{ij} - 2 \sum_{i=1}^n w_i \tilde{c}_{i0}, \quad (3.34)$$

în care:

- $\tilde{\sigma}_R^2$ - varianța erorii de estimare modelată;
- \tilde{c}_{00} - varianța modelată a valorilor măsurate;
- \tilde{c}_{ij} - covarianța modelată a valorilor măsurate;
- \tilde{c}_{i0} - covarianța modelată dintre valorile măsurate și punctul de estimare;
- w_i, w_j - ponderile acordate valorilor măsurate, a căror sumă este unitară.

Intervalul de încredere este utilizat pentru exprimarea valorii absolute a erorii de estimare pentru că ierarhizarea gradului de incertitudine al valorilor estimate nu este totdeauna suficientă pentru luarea unor decizii.

Un *interval de încredere* constă într-o *valoare minimă* și una *maximă* și în *probabilitatea* cu care valoarea reală se plasează în acest interval. Calculul intervalului de încredere pentru o valoare estimată prin kriging se realizează cu relația:

$$v^* \pm 2\tilde{\sigma}_R \quad (3.35)$$

care presupune că *erorile de estimare* au o *distribuție normală*, varianța de estimare modelată este o evaluare corectă a erorilor de estimare reale și că *probabilitatea ca valoarea reală să fie cuprinsă în acest interval este de 95%*.

Estimarea corectă a *intervalului de încredere* este determinată de alegerea modelului de variogramă și stabilirea corectă a distribuției erorilor de estimare care se presupune similară cu distribuția valorilor estimate.

Utilizarea varianței erorilor de estimare modelată (σ_R^2) pentru estimarea intervalului de încredere impune ca valoarea maximă a varianței modelate să fie o estimare corectă a varianței totale (efectul de scară). Diferența dintre varianța erorilor reale de estimare și cele estimate pe baza modelului probabilist este determinată de diferența dintre palierul modelului de variogramă și varianța totală reală. Deși valoarea palierului variogramei nu afectează valoarea estimată, ea este proporțională cu varianța estimării.

Interval de încredere pentru media globală. Estimarea mediei globale pentru o suprafață A cu caracteristici statistice uniforme poate fi suplimentată cu evaluarea intervalului de încredere asociat unei anumite erori de genul I , adecvată nivelului de precizie necesar într-o anumită etapă a cercetării.

Suprafața de calcul a mediei globale poate fi cea a unui bloc de exploatare dintr-un zăcământ de plumb sau zona contaminată cu petrol dintr-un acvifer freatic în care este necesară evaluarea grosimii medii a stratului de hidrocarburi acumulat la partea superioară a acviferului.

În fiecare din aceste cazuri asocierea valorilor medii ale variabilei pe suprafața studiată cu un interval de încredere aduce un plus de rigoare estimării.

Pentru calculul mediei globale există un singur set de date. În acest context evaluarea intervalului de încredere care se fundamentează pe ideea de repetabilitate se justifică totuși prin faptul că pot exista suprafețe similare din punct de vedere statistic și că pe aceeași suprafață s-ar putea obține și alte seturi de date.

Noțiunea de interval de încredere al mediei globale poate fi deci interpretată ca o cuantificare a fluctuațiilor în estimarea mediei globale de la un set de date la altul prelevate de pe aceeași suprafață sau de la o regiune la alta, similare din punct de vedere statistic.

Calculul intervalului de încredere al mediei globale se bazează pe varianța erorilor de estimare calculată cu formula:

$$\tilde{\sigma}_R^2 = \tilde{c}_{AA} + \sum_{i=1}^n \sum_{j=1}^n w_i w_j \tilde{c}_{ij} - 2 \sum_{i=1}^n w_i \tilde{c}_{iA}, \quad (3.36)$$

în care:

A - suprafața pe care se estimează media globală;

\tilde{c}_{AA} - media covarianțelor modelate pentru întreaga suprafață studiată,

\tilde{c}_{ij} - covarianța între toate punctele probate;

\bar{c}_{iA} - media covarianțelor dintre valorile din punctele probate și cele de discretizare.

Calculul varianței erorilor de estimare prin intermediul formulei (3.36) presupune că suma ponderilor (w_i) este unitară, variabila a cărei medie se calculează este staționară, iar modelul de covarianță este corect ales.

Evaluarea intervalului de încredere pentru media globală pe baza relațiilor (3.35) și (3.36) permite o estimare a fluctuațiilor acesteia pe baza modelului funcției aleatoare care de cele mai multe ori conduce la o supraevaluare a acestui interval. Experimental se constată că varianța erorii de estimare a unor alte seturi de date realizate pe aceeași suprafață este mai mică decât cea prognozată pe baza modelului funcției aleatoare. Din păcate, de cele mai multe ori nu avem acces la o nouă probare pe suprafețele studiate și deci nici posibilitatea de a corecta varianța erorii de estimare.

Interval de încredere local. Evaluarea intervalului de încredere pentru media globală presupune o configurație spațială similară pentru întreaga suprafață studiată, dar în anumite circumstanțe particularitățile locale trebuie luate în considerare.

Dacă în general distribuția erorilor este simetrică acest lucru nu este valabil în toate zonele suprafeței studiate. Din acest motiv, în zonele cu valori mici crește probabilitatea supraevaluărilor, iar în cele cu valori mari a subevaluărilor. La nivel local presupunerea normalității distribuției erorilor este acceptabilă doar în cazul anumitor proprietăți geometrice, cum ar fi grosimea unui strat de cărbune într-un zăcământ stratiform, adâncimea unui reper stratigrafic etc.

Chiar dacă acceptăm normalitatea distribuției erorilor, există dificultăți în calculul varianței locale. Dacă intenționăm să aplicăm formula (3.35) pentru evaluarea intervalului de încredere local, trebuie să ne asigurăm că modelul de variogramă pe care îl utilizăm este reprezentativ pentru caracteristicile spațiale ale zonei respective.

Cea mai simplă variantă pentru soluționarea acestor probleme este acceptarea ipotezei că modelul de variogramă este același pentru întreaga suprafață, diferența de la o zonă la alta fiind dată numai de parametrii acestuia și în special de palier.

Se definește un model de variogramă relativă, cu palierul unitar, a cărei formă descrie particularitățile spațiale ale variabilei și care se utilizează pentru generarea ecuațiilor sistemului de kriging. În mod practic, modelul variogramei relative se obține prin divizarea coeficienților modelului variogramei absolute prin valoarea palierului.

Varianța erorii de estimare care se obține este relativă la varianța locală și se calculează cu relația:

$$\tilde{\sigma}_R^2 = \tilde{c}_{00} + \sum_{i=1}^n \sum_{j=1}^n w_i w_j \tilde{c}_{ij} - 2 \sum_{i=1}^n w_i \tilde{c}_{i0}, \quad (3.37)$$

în care:

\tilde{c}_{00} - varianța valorilor;

\tilde{c}_{ij} - covarianța între punctele p_i și p_j ;

\tilde{c}_{i0} - covarianța între punctul p_i și punctul p_0 .

Pentru evaluarea varianței locale a erorilor se corectează valoarea obținută din relația (3.37) cu un estimator al varianței locale (σ^*)²:

$$\tilde{\sigma}_R^2 = (\sigma^*)^2 \cdot \left[\tilde{c}_{AA} + \sum_{i=1}^n \sum_{j=1}^n w_i w_j \tilde{c}_{ij} - 2 \sum_{i=1}^n w_i \tilde{c}_{iA} \right], \quad (3.38)$$

în care covarianțele se obțin prin diferența dintre unu și variogramele relative corespunzătoare.

A26 Distribuția spațială a erorilor de estimare

Să se calculeze erorile de estimare a distribuției spațiale pentru sarcina piezometrică a acviferului freatic din zona stației de tratare Ogrezeni (**Fig.3.9**).

Rezolvare:

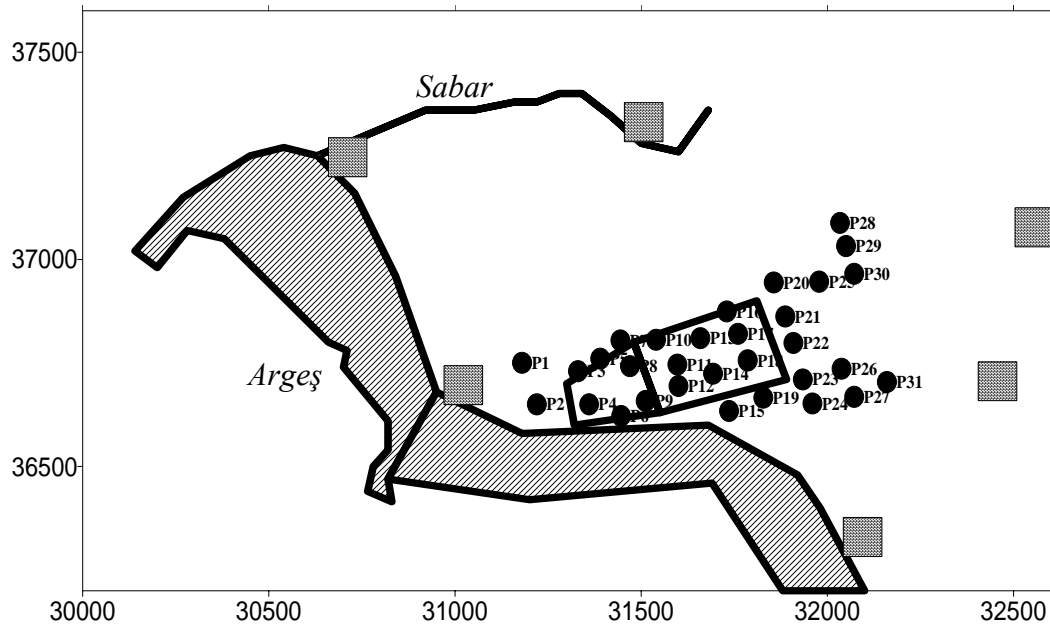


Fig.3.9. Rețeaua de monitorizare a nivelului piezometric

■ Stație de monitoring zonal ● Stație de monitoring local

Pentru evaluarea monitorizării nivelului piezometric s-a proiectat o rețea locală de 31 de piezometre amplasate pe aliniamente transversale pe valea Argeșului completată cu una zonală de 6 piezometre. Rețeaua proiectată nu a ținut seama de construcțiile din zonă și de posibilele surse de infiltrații ale rețelelor de conducte și bazinelor. Distanța rezultată între piezometre este cuprinsă între 50 și 250 m (**Fig.3.9**).

Sunt proiectate și două mire hidrometrice în zona stației de tratare.

Analiza variografică a nivelului piezometric al acviferului de mică adâncime a condus la următoarele rezultate:

- structura spațială a suprafeței pezometrice prezintă o tendință zonală orientată NV-SE;
- suprafața piezometrică are o *structură anizotropă* cu raportul de anizotropie $r_a = 1/3$ cu *direcția de continuitate maximă* orientată N35°E;
- modelul variogramei este *sferic* cu parametrii:
 - *efectul de pepită* = 0,00;
 - *palierul* = 25 m²;
 - *razele de influență*:
 - $r_1 = 2100$ m (pe direcția N35°E);
 - $r_2 = 700$ m (pe direcția N55°V).

Validarea

modelului de variogramă s-a realizat prin *kriging punctual universal* și rezultatele validării sunt prezentate în **tabelul 3.1**.

Analiza statistică a diferențelor între valorile măsurate și cele calculate ale cotei nivelului hidrostatic indică o repartiție cvasisimetrică a acestora (**Fig.310.**) și o valoare medie a erorilor de 0,03.

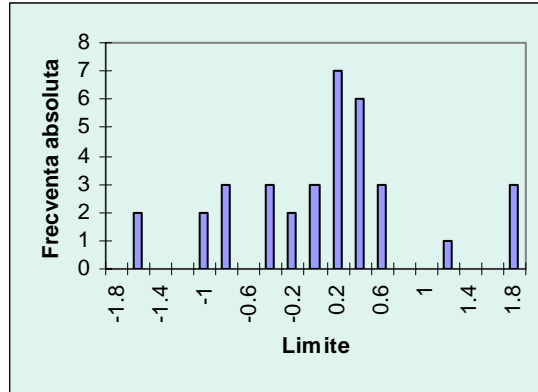


Fig.3.10. Histograma erorilor validării

Tabelul 3.1. Rezultatele validării modelului de variogramă

x	y	Cota NH măsurat	Cota NH calculat	Dif.	x	y	Cota NH măsurat	Cota NH calculat	Dif
30900	36850	101.9	100.17	-1.7	31572	36775	96.75	96.789	0.0
31440	36780	98.15	96.524	-1.6	30960	37380	101	101.1	0.1
31748	36743	96.61	95.487	-1.1	31910	36700	95.15	95.298	0.1
31944	36882	97.75	96.649	-1.1	31570	36800	96.9	97.051	0.2
31420	37370	101	100.13	-0.9	32180	37020	97.06	97.256	0.2
32334	36652	95.21	94.397	-0.8	30850	36970	101.58	101.79	0.2
30580	37280	102.4	101.6	-0.8	32426	37010	96.72	96.939	0.2
32221	36930	97.15	96.562	-0.6	30260	37230	102.05	102.28	0.2
30720	37180	101.63	101.2	-0.4	31860	36570	93.34	93.631	0.3
31370	36570	95.69	95.263	-0.4	32000	36980	97.3	97.639	0.3
32140	36610	93.94	93.546	-0.4	30650	37240	101.31	101.7	0.4
31600	36660	95.6	95.347	-0.3	31450	36670	96.2	96.637	0.4
32204	37040	97.26	97.095	-0.2	30700	37280	101	101.46	0.5
32443	36856	96.36	96.287	-0.1	32058	36780	95.53	96.057	0.5
31220	37400	101	101	0.0	32210	36610	93.32	94.344	1.0
31630	36670	95.47	95.476	0.0	31220	36620	95.58	97.685	2.1
30460	37270	102.25	102.27	0.0	31690	36550	92.2	94.516	2.3
31572	36775	96.75	96.789	0.0	31690	36550	92.2	94.516	2.3

Distribuția abaterii standard de evaluare s-a calculat prin *kriging punctual universal*, cu o tendință zonală modelată printr-o *suprafață polinomială de gradul al doilea*, cu o *abatere standard la calare de 10%*. Rețeaua de calcul a fost pătratică, formată din 1326 de celule ($n_x = 51$; $n_y = 21$).

Abaterea standard obținută are valori cuprinse între 0,528 și 4,76 m. Cu cele 1326 de valori ale abaterii standard s-au calculat *erorile de estimare relative* cu:

$$\varepsilon_r = \frac{2 \cdot KSD}{CotaNH} \cdot 100 [\%] \quad (3.39)$$

în care

KSD - abaterea standard de kriging ;

CotaNH - cota nivelului hidrostatic.

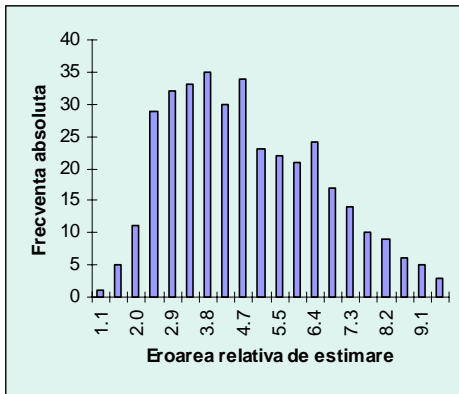


Fig.3.11. Histograma erorilor relative de estimare

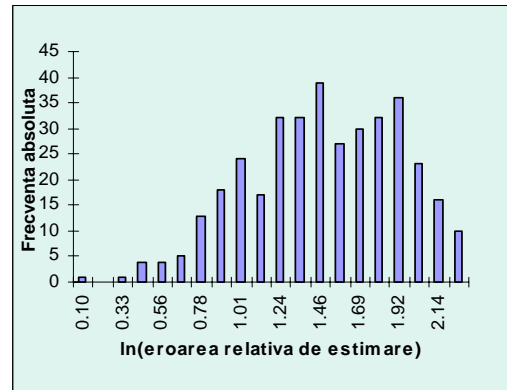


Fig.3.12. Histograma erorilor relative de estimare logaritmice

Analiza statistică a erorilor relative de estimare indică o selecție asimetrică cu repartiție lognormală (**Fig.3.11** și **Fig.3.12**) care conduce la o subestimare a mediei acestora cu 9,5%.

Valorile erorilor relative de estimare a cotelor nivelului piezometric sunt cuprinse în intervalul 1,1%...9,3% cu $\varepsilon(\alpha) = +/-0,19\%$. Rezultă că în toată zona cercetată erorile de estimare ale suprafeței piezometrice nu depășesc, pentru regimul natural de curgere, 10% din valoarea cotelor măsurate.

Distribuția spațială a acestor erori este vizualizată sub forma unei hărți cu izolinii (**Fig.3.13**) din care se remarcă:

- plasarea arealului cercetat în zona erorilor relative de estimare mai mici de 5,5% (zona centrală din **Fig.3.13**);
- plasarea platformei realizate din umplutură în zona erorilor relative mai mici de 3%;
- conturarea unei zone cu erori mai mari (>4%) în partea centrală a ariei cercetate determinată probabil de alimentări active din infiltrații, favorizate de permeabilitatea crescută a depozitelor de umplutură.

COMENTARIU

Dacă operăm fără greșală până în această etapă de prelucrare, erorile pe care le obținem prin calcul sunt rezultatul gradului incomplet de cunoaștere al variabilității spațiale a caracteristicii studiate (în cazul aplicației, cota nivelului piezometric). Erorile de estimare pe care le-am calculat nu trebuie să fie determinate de greșelile pe care le-am comis noi în etapa de estimare.

Orice eroare datorată prelucrărilor noastre poate fi eliminată prin reluarea etapelor de prelucrare. Pe lângă neatenție, sursa principală a erorilor este alegerea unui model de variogramă neadecvat.

Indiferent de rezultatele testării modelului de variogramă, eu nu definitivez nici o estimare spațială până nu am utilizat pentru aceeași variabilă cel puțin două modele de variogramă!

Reducerea erorilor de estimare determinate de insuficienta cunoaștere a structurii spațiale studiate poate fi realizată doar prin completarea informației primare cu măsurători suplimentare

3.2. REDUCEREA ERORILOR

Toate informațiile obținute din rețelele de cercetare au ca obiectiv *reducerea incertitudinilor* în legătură cu distribuția spațială a parametrilor măsurați (ex.: în *aer*: SO₂, NO₂, pulberi sedimentabile etc.; în *apele de suprafață*: NO₃, NH₄, pH, temperatură etc.; în *sol*: conținut în substanțe organice sau substanțe poluante, etc.; în *subsol*: conținutul în metale al filoanelor, conținutul în substanțe poluante al apelor subterane etc.). Reducerea incertitudinilor implică metodologic teoria *bayesiană* și *entropia* proceselor aleatoare, motiv pentru care criteriile care stau la baza proiectării rețelelor de cercetare sunt: *minimizarea entropiei* și *reducerea erorilor de interpolare*.

Utilizarea entropiei este singurul mijloc de a cuantifica incertitudinea globală relativă la variabilele măsurate în rețele de monitoring pe baza legilor de distribuție identificate experimental (prin metode de calare a modelelor teoretice).

În condițiile utilizării entropiei criteriul de optimizare a rețelelor de monitoring este minimizarea acesteia.

Entropia

Entropia reprezintă reducerea incertitudinii unui *eveniment* care se produce cu *probabilitatea p*:

$$H = -\log(p) \quad (3.40)$$

Entropia este *aditivă* când este aplicată intersecției a două *evenimente independente*:

$$H = p(-\log p) + (1-p)[- \log(1-p)] \quad (3.41)$$

Cele două evenimente pot fi două distribuții punctuale cu probabilitățile p și $(1-p)$ și această definiție poate fi extinsă la distribuția mai multor puncte (ex.: punctele de probare ale rețelei de cercetare).

Entropia este definită strict pozitiv pentru un set de valori discrete de volum finit. Ea este maximă pentru o distribuție uniformă a acestora (cu aceeași probabilitate și lege de distribuție). Considerând o serie de valori infinit apropiate, utilizându-se *integrala* în locul *sumei* și *funcția densității de probabilitate* în locul *probabilităților*, se definește o entropie analoagă sub forma:

$$H(V) = E[-\log f(V)] \quad (3.42)$$

Analogia cu mediul discret nu este satisfăcătoare deoarece nu este invariantă la transformările realizate asupra variabilei aleatoare V , *funcția densității de probabilitate* fiind exprimată în *unități de probabilitate* raportate la *unități de V*. Pentru uniformizare dimensională este propusă (Jaynes, 1968) forma:

$$H(V) = E[-\log f(V)/m(V)] \quad (3.43)$$

în care $m(V)$ este o măsură a "ignoranței complete" asupra variabilei V . Utilizată în această formă, cu toate ambiguitățile introduse de alegerea lui $m(V)$ entropia își păstrează proprietatea de *aditivitate* dar și-o pierde pe cea de *pozitivitate*. Ca o măsură

a incertitudinii în mediul continuu este utilizată valoarea absolută a entropiei calculate cu formula lui Jaynes.

Entropia rețelei de monitoring

Abordarea studiului *entropiei rețelei de monitoring* pornește de la ipoteza că datele furnizate de aceasta pentru cartografierea unor parametri sunt dintr-un spațiu *aleator* generat de un proces natural.

Pentru a simplifica reprezentarea ne putem imagina acest spațiu aleator bidimensional, în el fiind dispersate valorile parametrului măsurate în stațiile rețelei de monitoring.

Toate valorile măsurate se consideră *reprezentative* pentru fenomenul monitorizat, dar nu toate locațiile investigate prin măsurători candidează pentru funcția de *stații permanente în rețeaua de monitoring*. Densitatea stațiilor de măsurare este limitată din considerente practice (costul execuției și al măsurătorilor executate). Minimizarea entropiei vizează stabilirea numărului minim de stații permanente din rețea.

Fie X vectorul *aleator* ale cărui coordonate reprezintă valorile câmpului spațial în diferitele puncte ale rețelei de monitoring, ordonate după un criteriu rezonabil.

Stațiile stabile ale rețelei de monitoring vor fi selectate dintre aceste puncte disponibile iar problema proiectantului este identificarea acelor care trebuie măsurate.

Pentru această operațiune partiția vectorului X poate fi reprezentată sub forma:

$$X = (U, G) \quad (3.44)$$

în care

U - reprezintă valorile variabilei în punctele nemonitorizate (stații nepermanente);

G - reprezintă valorile variabilei în puncte monitorizate permanent.

Pentru a simplifica notațiile se va nota cu $f(U, G)$ funcția densității de probabilitate condiționată a variabilei $X = (U, G)$ și cu $m(U, G)$ măsura *completei ignoranțe*. Presupunând că $m(G)$ este specificat, atunci :

$$m(U|G) = m(U, G)/m(G) \quad (3.45)$$

Pentru un anumit proces stocastic, incertitudinea asupra rețelei de monitoring se exprimă prin entropia variabilei X :

$$H(X) = H(U|G) \quad (3.46)$$

care se descompune în :

$$H(U, G) = H(U|G) + H(G) \quad (3.47)$$

în care

$$H(U|G) = E[-\log f(U|G)/m(U|G)] \quad (3.48)$$

Pentru a interpreta această decompoziție, presupunem că valorile G au fost obținute și în aceste condiții incertitudinea asupra valorilor U este :

$$H(U) = E[-\log f(U|G)/m(U|G)|G] \quad (3.49)$$

Deoarece G nu este cert, incertitudinea relativă la U este primul termen în decompoziția entropiei $H(X) = H(U, G)$:

$$H(U|G) = E\{E[-\log f(U|G)/m(U|G)|G]\} \quad (3.50)$$

deci

$$H(X) = H(U, G) = E\{E[-\log f(U|G)/m(U|G)|G]\} + E[-\log f(G)/m(G)] \quad (3.51)$$

Strategia proiectării derivă din decompoziția entropiei $H(X)$. Minimizarea incertitudinii se realizează prin minimizarea termenului $H(U|G)$ sau maximizarea termenului $H(G)$.

Maximizarea termenului $H(G)$, care reprezintă incertitudinea a priori asupra lui G , poate fi eliminată prin monitorizare. Este atunci ușor de intuit că acest termen poate fi maximizat printr-o partiție corespunzătoare a lui X , maximizându-se astfel beneficiul măsurătorilor executate.

Se demonstrează că:

$$H(U|G) = E[-\log f(U|G)/m(U|G)] < E[-\log f(U)/m(U)] = H(U) \quad (3.52)$$

În aceste condiții, obținând G prin măsurători niciodată nu va avea loc o creștere a incertitudinii asupra partiției U . Presupunând $m(U|G) = m(U)$ rezultă:

$$H(U|G) = H(U) - I(U; G) \quad (3.53)$$

în care informația mutuală în U și G este strict pozitivă, simetrică și definită de relația:

$$I(U; G) = E\{\log[f(U, g)/(f(U)f(G))]\} \quad (3.54)$$

Obiectivul minimizării entropiei se realizează în contextul instalării unei noi rețele sau al extinderii uneia existente. În ambele cazuri, efectul unui număr restrâns de puncte măsurate suplimentar, repartizate pe întreaga suprafață monitorizată, este minim asupra reducerii entropiei. În același timp potențialul de inferență al partiției U din G devine de critică importanță atâta timp cât rețeaua de monitoring este utilizată repetat în acest scop.

Criteriul postulat pentru reducerea incertitudinii asupra lui U obținută prin G este maximizarea informației mutuale I care ignoră consecințele entropiilor $H(U)$ și $H(G)$ în partiționarea vectorului X :

$$H(U) - H(U|G) = I(U; G) = I \quad (3.55)$$

Posibilitatea includerii în U a valorilor cu incertitudine mare este admisă pe baza posibilității comode de a fi deduse din G . Această aproximare este strâns legată de teoria transmiterii informației a lui *Shannon*.

Situația cu care ne confruntăm deseori în monitorizarea unui proces este cea a unei rețele de observație existente, interesul fiind acela de a elimina stațiile de observație care aduc o cantitate redusă de *informație utilă*.

Deoarece rețelele sunt constituite, în general, dintr-un număr redus de puncte, incertitudinea conținută în $H(G)$ reprezintă o fracțiune semnificativă din cea totală ($H(X)$) și maximizând $H(G)$ minimizăm $H(U|G)$.

În practică, modelele stocastice necesare pentru calculul diferitelor probabilități vor fi ele însele incerte. Multe teorii statistice sunt elaborate pentru evaluarea și descrierea acestei incertitudini sub forma unui vector " θ ". Se poate încorpora în mod direct această incertitudine a modelului în incertitudinea asupra distribuției procesului analizat sub forma:

$$H(U, G, \theta) = H(U, G | \theta) + H(\theta) \quad (3.56)$$

$H(\theta)$ - incertitudinea asupra parametrilor modelului stocastic utilizat;

$H(U, G, \theta)$ - incertitudinea totală a variabilei X .

În studiul rețelilor de monitoring se pleacă de cele mai multe ori de la un set de date minim. Toate probabilitățile și entropiile calculate sunt condiționate de reprezentativitatea și corectitudinea acestuia.

Entropiile instrumentale

Toate formele de erori de măsurare trebuie luate în considerare la proiectarea rețelilor de monitoring deși în mod surprinzător sunt rețele care nu dispun de măsurători de verificare (duble).

Incetitudinea asupra variabilei $X = (U, G)$ este cea care interesează în principal iar G , presupusă a fi furnizată de rețea, este importantă pentru definirea procesului optim de măsurare. Operațiunea de măsurare induce erori iar măsurătorile duble sunt utilizate pentru evaluarea mărimii acestora.

Fie D vectorul tuturor măsurătorilor disponibile executate în punctele care formează variabila G . Reducerea incertitudinii datorată vectorului D este:

$$H(U, G, \theta) - H(U, G, \theta | D) = H(G) - H(G | D) \quad (3.57)$$

cu măsura *ignoranței complete*:

$$m(U, G, \theta | D) = m(U, \theta | G)m(G | D) \quad (3.58)$$

Atât timp cât $H(U, G, \theta)$ este constantă și independentă de alegerea procesului de măsurare a variabilei G și de modul de partiționare (U, G) , se deduce că $H(U, G, \theta | D)$ poate fi minimizată prin combinația modului de alegere a procesului de măsurare (G) și de partiționare a variabilei X .

În general $f(U, \theta | G, D) = f(U, \theta | G)$ și dacă $m(U, G, \theta | D) = m(U, \theta | G)m(G | D)$ rezultă că:

$$H(U, G, \theta | D) = H(U, \theta | G) + H(G | D) \quad (3.59)$$

Alegerea metodei de măsurare afectează numai termenul al doilea ($H(G|D)$) în timp ce *alegerea rețelei* (partiția variabilei X) afectează ambii termeni. În mod ideal alegerea rețelei și procesului de obținere a datelor trebuie să se facă simultan. Procesul de *măsurare* trebuie să includă desemnarea *laboratoarelor de analiză*. Calitatea datelor poate depinde nu numai de laboratorul în care se execută analiza ci și de durata transportului probei până la laborator. Cu alte cuvinte, *mărimea erorilor* poate depinde de *distanța dintre stația monitorizată și poziția laboratorului*.

Dacă $m(D|G) = m(G|D)m(D)/m(G)$ atunci

$$H(G|D) = H(G) - [H(D) - H(D|G)] \quad (3.60)$$

relație care exprimă faptul că o bună cunoaștere a variabilei G reduce substanțial incertitudinea asupra distribuției variabilei X .

Normalitatea multivariată

Transformate în mod adecvat multe serii de date au distribuții care pot fi approximate printr-o *distribuție gaussiană*, distribuție adecvată prelucrărilor cu modele geostatistice.

În contextul noțiunilor introduse în secțiunea anterioară, variabila X este vectorul valorilor parametrilor mășurați în stațiile unei rețele de monitoring la un moment dat. Coloanele matricii de date $D(X_1, X_2, \dots, X_n)$ sunt presupuse a priori a fi independente.

O altă ipoteză plauzibilă este aceea că $\{X_i\}$ are o distribuție normală multivariată cu un vector mediu μ și matricea de covarianță Σ , distribuție care poate fi exprimată simbolic:

$$X_i | \mu, \Sigma \sim \text{ind } N_p(\mu, \Sigma) \quad (3.61)$$

p - numărul de puncte de observație din rețea;

n - numărul de momente la care s-au executat măsurători ($i = 1 \dots n$)

Problema care se rezolvă în ipoteza *distribuției normale multivariate* este cea a partiției selecției X în U și G astfel încât prin monitorizarea partiției G să se reducă la maximum entropia $H(X)$.

Pentru descompunerea *entropiei totale* matricea de covarianță se reparametrizează ca $(\Sigma_{GG}, \Sigma_{U|G}, \tau)$ conform relațiilor:

$$\Sigma_{GG} = \text{cov}(G, G)^T = E[G - E(G)][G - E(G)]^T \quad (3.62)$$

$$\Sigma_{U|G} = \Sigma_{UU} - \Sigma_{UG} \cdot \Sigma_{GG}^{-1} \cdot \Sigma_{GU} \quad (3.63)$$

Matricea $\tau = \Sigma_{U|G} \cdot \Sigma_{GG}^{-1}$ este panta predictorului linear optimal al partiției U prin G (adică $E(U) + \tau[G - E(G)]$).

Această transformare este realizată prin decompoziția Bartlett $\Sigma = T \Delta T^T$, în care

$$\Delta = \begin{bmatrix} \sum_{U|G} & 0 \\ 0 & \sum_{GG} \end{bmatrix} \quad T = \begin{bmatrix} I & \tau \\ 0 & I \end{bmatrix} \quad (3.64)$$

$$\Sigma^I = (T^T)^{-1} \Delta^{-1} T^{-1} \quad (3.65)$$

$$\Delta^{-1} = \begin{bmatrix} \sum_{U|G}^{-1} & 0 \\ 0 & \sum_{GG}^{-1} \end{bmatrix} \quad T^{-1} = \begin{bmatrix} I & -\tau \\ 0 & I \end{bmatrix} \quad (3.66)$$

Pentru un set de date $d(X_1, X_2, \dots, X_n)$ descompunerea entropiei se face în trei componente:

$$E[-\log f(X, |d)/m(X, |d)] = HU + HMODEL + HG \quad (3.67)$$

- HU reprezintă incertitudinea reziduală apriori a partiției U după utilizarea partiției G pentru evaluarea valorilor partiției U pe baza modelului liniar:

$$HU = E[-\log f(U|G), \mu, \Sigma | d] \quad (3.68)$$

- HG reprezintă incertitudinea asupra partiției G eliminată prin monitorizare și necesar de maximizat printr-o partiție optimă a variabilei X :

$$HG = E[-\log f(G|d)|d] \quad (3.69)$$

- $HMODEL$ reprezintă incertitudinea reziduală asupra partiției U , mediei și matricii de covarianță corespunzătoare unei anumite partiții G măsurate:

$$\begin{aligned} HMODEL = E[-\log f(\mu_U | \mu_G, G, \Sigma, d) | d] + E[-\log f(\mu_G | G, \Sigma, d) | d] + \\ + E[-\log f(\Sigma | G, \Psi, m, d) / m(\Sigma) | d] \end{aligned} \quad (3.70)$$

în care Ψ și m sunt parametri ai distribuției Wishart (Anderson, 1984). Optimizarea rețelei se fundamentează pe minimizarea componentelor HU și $HMODEL$.

Reducerea erorii de estimare zonală se poate realiza doar prin îndeșirea punctelor de observație. Evaluarea efectului amplasării unui nou punct de observație este posibilă prin intermediul *varianței erorii de estimare* care depinde numai de *modelul de variogramă* și de *distanța* dintre punctul în care se face estimarea și punctele de observație (nu depinde de valorile măsurate !!!).

Câștigul de precizie ($CP(p_0)$) asociat unui punct de estimare (p_0) prin introducerea unui *punct fictiv* în zona lui de influență se estimează cu relația:

$$CP(p_0) = \frac{(\tilde{\sigma}_R^2)_f - (\tilde{\sigma}_R^2)_i}{(\tilde{\sigma}_R^2)_f} \quad (3.71)$$

în care

$(\tilde{\sigma}_R^2)_f$ este varianța erorii de estimare după introducerea *punctului fictiv*;

$(\tilde{\sigma}_R^2)_i$ - varianța erorii de estimare inițială.

Eficiența amplasării noilor puncte de observație se apreciază după valoarea *câștigului de precizie* obținut.

Proiectarea rețelei de monitoring

Trei etape principale sunt implicate de proiectarea unei rețele de monitoring:

- *inițializarea rețelei de monitoring*;
- *extinderea rețelei de monitoring*;
- *optimizarea rețelei de monitoring*.

Aceste etape vor fi analizate doar din punctul de vedere al contribuției lor la *reducerea erorilor de estimare* determinate de distribuția spațială a locațiilor stațiilor de observație din rețeaua de monitoring.

Inițializarea rețelei de monitoring se realizează în condițiile în care nu au fost colectate nici un fel de date într-un mod sistematic în suprafața ce urmează a fi monitorizată. Informațiile disponibile sunt în această fază limitate ca volum și în mare măsură subiective, astfel încât orice plan rațional este aparent exclus.

Cea mai indicată alegere în aceste circumstanțe este o rețea de monitoring uniformă (cu amplasarea stațiilor de monitorizare într-o rețea pătrată). În practică, realizarea unei astfel de rețele este împiedicată de gradul de accesibilitate la punctele proiectate în teren, asigurarea securității stațiilor de monitorizare, dorința de a minimiza costul execuției, dorința de a realiza un număr maxim de stații dintr-un buget fix.

Pentru monitorizarea parametrilor ambientali, de cele mai multe ori rețelele de monitorizare s-au corelat mai mult cu densitatea populației decât cu factorii ambientali. Când cunoașterea distribuției spațiale este vagă, o bază obiectivă și rațională pentru proiectarea unei rețele de monitoring necesită anumite date colectate înaintea realizării unei *rețele permanente*. Numai dacă nu se pot obține prin *teledetecție*, pentru datele minime necesare se poate realiza un sistem de *stații temporare*.

Trebuie făcută distincția între locațiile care sunt potențiale stații permanente și care sunt monitorizate temporar pentru a obține informații statistice pentru procesul de proiectare.

Corelarea unei cunoașteri inițiale și a unor date obținute din rețele temporare cu schema bayesiană a reducerii entropiei aduce o importantă contribuție la descrierea statistică adecvată a fenomenelor regionale ambientale, cu *cea mai scurtă perioadă de funcționare a rețelelor temporare*.

Două alternative ale criteriului *entropiei* sunt aplicabile:

- Maximizarea informației mutuale $I(U, G)$ (3.54).

Aplicarea acestui criteriu plasează centrul de greutate asupra informațiilor obținute din *rețeaua temporară*. Sunt necesare date în toate stațiile din partițiile U și G în scopul stabilirii funcțiilor de frecvență $f(U)$ și $f(U|G)$. *Rețeaua temporară* de monitoring trebuie să conțină un număr suficient de stații pentru a realiza o bună estimare a partiției U chiar dacă multe vor fi numai temporare. Valoarea maximă

$I(U;G)$ va crește la început proporțional cu numărul stațiilor până când numărul de stații din G îl depășește pe cel din U . Acest criteriu este eficient pentru rețele cu un număr de stații (g) mult mai mic decât cel al stațiilor temporare (p).

- Minimizarea incertitudinii $HU + HMODEL$ (3.68-3.70).

Acest criteriu plasează pe ultimul loc rolul informațiilor obținute din rețeaua temporară utilizată pentru proiectarea rețelei deoarece echivalentul lui este maximizarea parametrului $|\Psi_{GG}|$. El trebuie evaluat numai pentru un set de locații selectate dintre stațiile care candidează la statutul de *stații permanente*. Stațiile temporare nu contribuie la procesul de proiectare deși teoria ia în considerare în mod explicit predicția valorilor în stațiile temporare. Pentru compararea eficienței unor rețele de dimensiuni diferite trebuie calculați parametri suplimentari pe baza unei rețele extinse care să cuprindă și stațiile temporare. Pe măsură ce numărul stațiilor monitorizate crește, valoarea minimă a incertitudinii $HU + HMODEL$ descrește în timp ce valoarea maximă HG crește. Alegerea *numărului optim de stații* în rețeaua de monitoring se face prin evaluarea efectului introducerii fiecărei locații asupra reducerii incertitudinii $HU + HMODEL$.

Extinderea unei rețele existente este similară cu inițializarea unei rețele, diferența constând în faptul că primul grup de stații necesare pentru obținerea unor informații preliminare sunt *predeterminate*.

Criteriul favorizat este cel al minimizării incertitudinii $HU + HMODEL$ (în defavoarea celui de minimizare a informației mutuale $I(U;G)$). În prelucrare trebuie să se țină seama că va fi un puternic dezechilibru între calitatea informației din *rețeaua preexistentă* și cea nouă. Informațiile din rețeaua existentă au o mai lungă *istorie* decât cele obținute din noile locații.

Optimizarea unei rețele existente este abordată în același mod ca inițializarea ei dar pe baza unor *informații suplimentare* obținute din rețeaua respectivă pe o anumită perioadă. Acest avantaj poate să dispară dacă distribuția locațiilor și programul de măsurători realizat nu sunt adecvate descrierii variabilității caracteristicilor monitorizate.

Metoda *punctului fictiv* operează în această etapă în două direcții:

- *Eliminarea locațiilor ineficiente din rețeaua de monitoring*. Pentru fiecare punct de observație din rețeaua de monitoring se evaluează *câștigul de precizie* (3.71) și se elimină cele care nu contribuie la reducerea erorilor de estimare în mod semnificativ.

- *Completarea rețelei de monitoring*. În zonele cu erori de estimare mai mari decât valoarea admisă se amplasează *locații fictive* și se calculează eficiența lor prin intermediul *câștigului de precizie* pe care îl determină.

Este frecventă utilizarea *hârșilor cu izolinii de câștig de precizie* realizate prin îndesirea rețelei existente printr-o rețea regulată de locații fictive. Selectarea locațiilor unde se execută stații suplimentare se face pe baza unui *câștig minim de precizie* care se alege în funcție de gradul de detaliu solicitat în estimările spațiale realizate.

A27. Optimizarea rețelei de monitoring

Să se optimizeze rețeaua de explorare hidrogeologică a acviferului sub presiune din culcușul stratului V de lignit, în interfluviul Motru-Jiu.

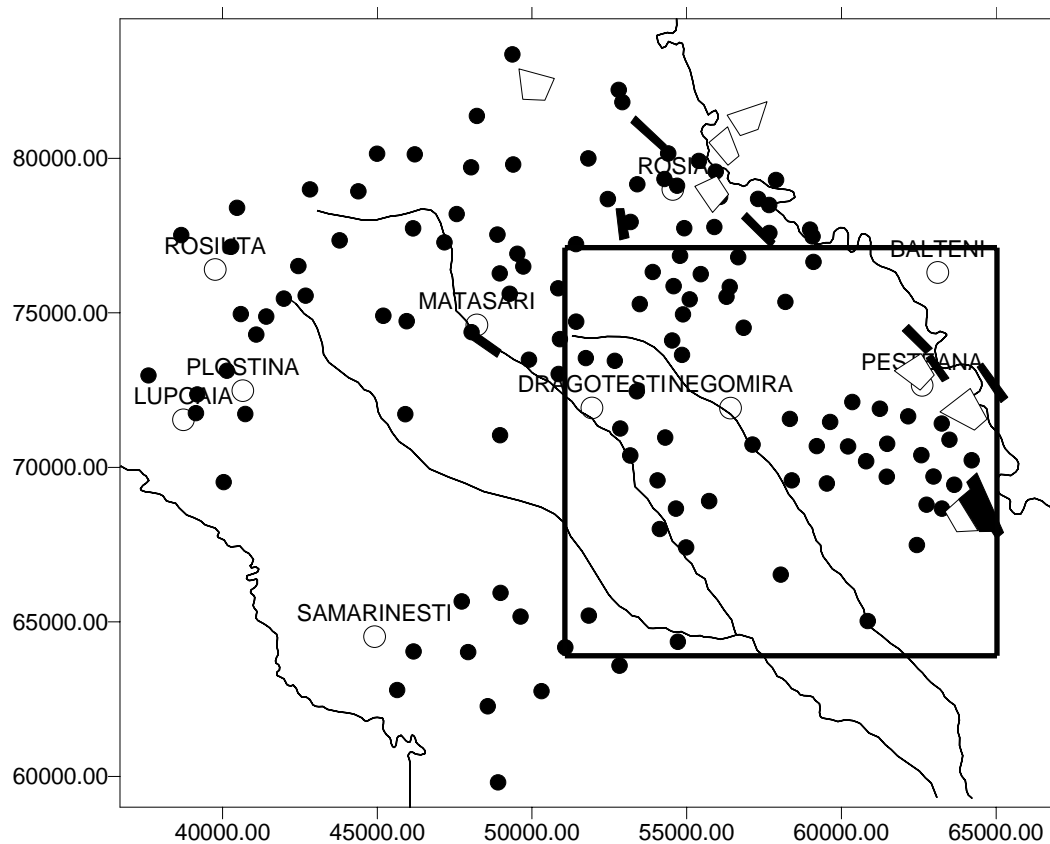


Fig.3.14. Amplasarea forajelor hidrogeologice de cercetare a orizonturilor acvifere

□ - sectorul selectat pentru aplicație.

Rezolvare:

Pentru cercetarea distribuției parametrilor hidrogeologici ai acviferelor din zona interfluviului Motru-Jiu (**Fig.3.14**) au fost realizate 129 de foraje hidrogeologice care au explorat diferite orizonturi acvifere și în care s-au realizat selectiv teste hidro dinamice.

Deoarece distribuția spațială a parametrilor hidrogeologici este diferită de la un parametru la altul optimizarea rețelei de cercetare se face având în vedere unul dintre aceștia. De obicei se alege parametrul cel mai important pentru cercetarea realizată.

În zona aleasă *transmisivitatea acviferului din culcușul stratului V* este un parametru important pentru stabilirea potențialului de debitare al acestui acvifer cu extindere regională.

Pentru acest acvifer se dispune de 39 de valori de transmisivitate, cuprinse între 10 și 130 m²/zi, cu distribuție lognormală și un coeficient de asimetrie de 1,48. După eliminarea unui număr de 7 valori considerate nereprezentative pentru selecția de date disponibilă, s-a normalizat distribuția celor 32 de valori rămase, prin logaritmare.

Variograma de suprafață calculată pentru valorile logaritmice indică o slabă anizotropie care în condițiile erorilor de determinare a distribuției spațiale a

transmisivităților (**Fig.3.15**) este neglijabilă. Modelul *variogramei omnidirectionale* utilizat pentru estimarea distribuției spațiale a transmisivității este de tip *sferic* cu:

- *Efectul de pepită*: $c_0 = 0,56$
- *Palierul*: $c = 3,00$
- *Raza de influență*: $r = 7000 \text{ m}$.

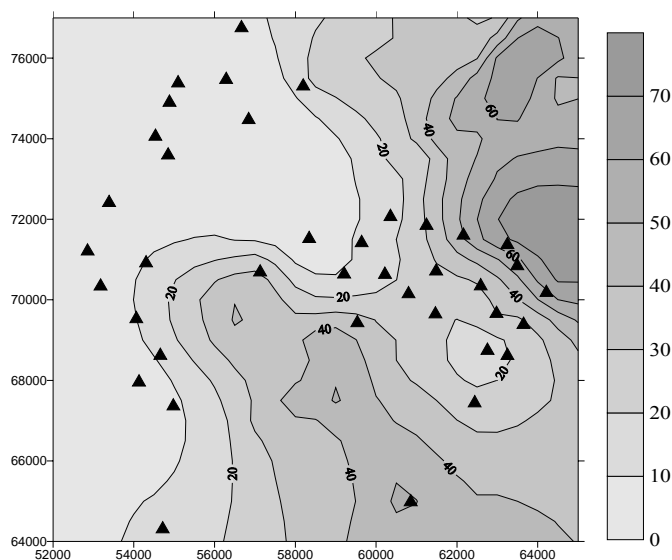


Fig.3.15. Distribuția transmisivității acviferului din culcușul stratului V calculată cu valori normalizate.

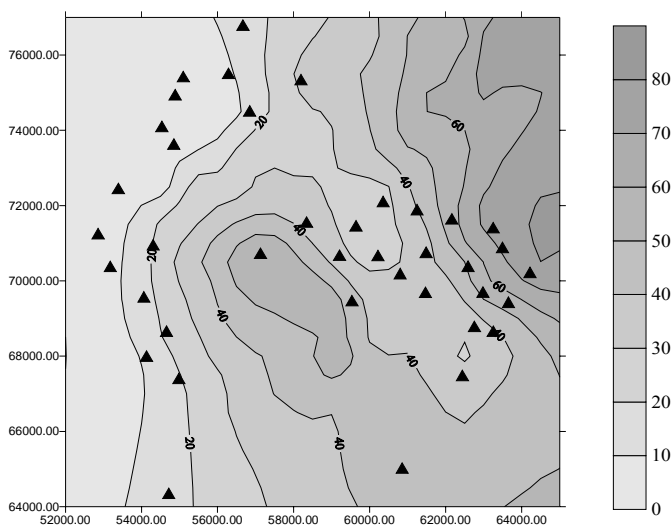


Fig.3.16. Distribuția transmisivității acviferului din culcușul stratului V calculată cu valori nenormalizate.

Normalizarea distribuției valorilor transmisivității permite evitarea supraestimării acesteia. Pentru ilustrarea efectului de supraestimare cauzat de asimetria distribuției datelor originale s-a realizat estimarea distribuției transmisivităților operându-se cu valorile originale, nelogaritmate (**Fig.3.16**).

Diferența dintre cele două estimări (**Fig.3.17**), cea *incorectă*, realizată cu valorile originale (cu distribuție logaritmică; **Fig.3.16**), și cea *corectă*, realizată cu valori normalizate (**Fig.3.15**), este semnificativă.

Diferența maximă dintre cele două estimări (datorată supraestimărilor) este de $38 \text{ m}^2/\text{zi}$, adică 50% din valoarea maximă estimată.

Evaluarea distribuției erorii de interpolare pentru estimarea distribuției transmisivității cu valori logaritmice (cea corectă, din **Fig.3.15**) s-a realizat prin *kriging zonal pe blocuri* rectangulare de $200 \times 300 \text{ m}$ cu 16 puncte de discretizare.

Harta conturală a abaterii standard calculată prin *kriging* (**Fig.3.18**) indică valori maxime localizate în zona nord vestică a interfluviului și acolo unde densitatea forajelor hidrogeologice este mai mică. În zonele de amplasare a forajelor în care a fost determinată transmisivitatea, *abaterea*

standard de estimare este mai mică de $0,7 \text{ m}^2/\text{zi}$ iar în zonele periferice ale perimetrului cercetat ajunge la valori maxime de $2 \text{ m}^2/\text{zi}$.

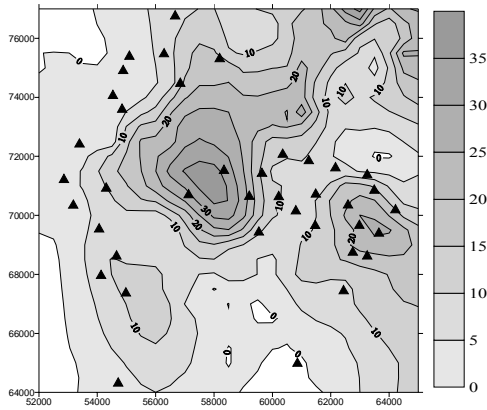


Fig.3.17. Distribuția supraestimărilor transmisivității acviferului din culcușul stratului V

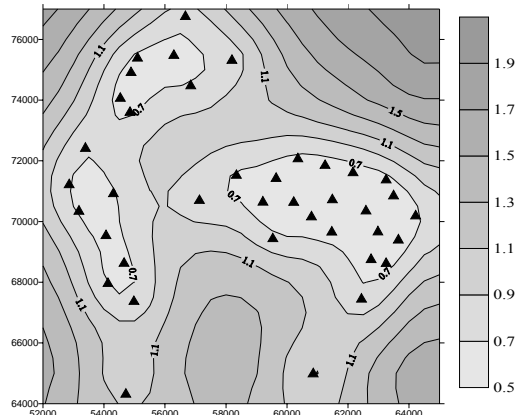


Fig.3.18. Distribuția abaterii standard de estimare a transmisivității acviferului din culcușul stratului V

Evaluarea efectului amplasării unei rețele regulate de puncte de observație fictive pe zona cercetată (**Fig.3.19**) se realizează cu metoda punctului fictiv și se exprimă prin câștigul de precizie.

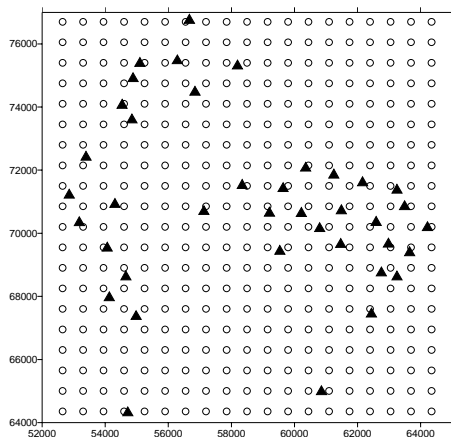


Fig.3.19. Distribuția punctelor fictive (○) și a forajelor hidrogeologice (▲) din perimetru.

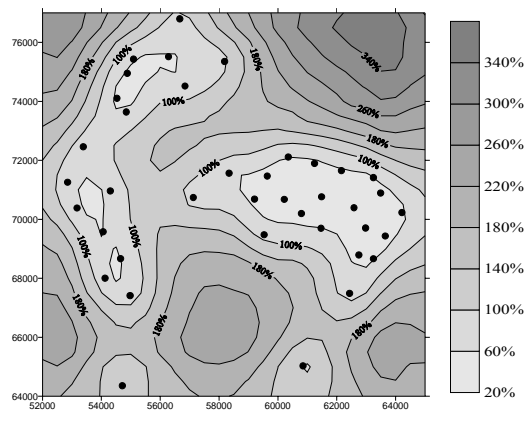


Fig.3.20 Distribuția câștigului de precizie determinat de introducerea celor 380 de puncte fictive

Pentru ilustrarea cantitativă a efectului, au fost amplasate 380 de puncte de observație fictive într-o rețea pătratică (19 coloane și 20 rânduri, cu parametrul rețelei 650 m).

Utilizându-se același model de variogramă (sferic, cu efect de pepită 0,56, palier 3 și rază de influență 7000 m) s-a obținut:

- un câștig maxim de precizie de 380% în zonele periferice ale zonei cercetate;
- un câștig minim de precizie de 20% în zonele de densitate maximă a forajelor hidrogeologice (**Fig.3.20**).

Harta cu distribuția câștigului de precizie (**Fig.3.20**) este utilizată pentru a alege zonele în care punctele fictive sunt considerate eficiente, adică determină un *câștig de precizie* semnificativ în estimarea distribuției transmisivității.

Astfel, dacă interesează distribuția transmisivității în zona nord-estică, este evident că amplasarea unor puncte de observație va aduce o creștere a preciziei de peste 300% și realizarea lor este eficientă. Amplasarea unor puncte suplimentare în zona de concentrare a forajelor hidrogeologice din partea centrală va aduce un câștig de precizie de numai 20% și eficiența lor este discutabilă.

Desigur că aceste *câștiguri* procentuale se pot transforma prin intermediul *intervalului de încredere* din *abatere standard* în *valori absolute* (3.35), care pot avea semnificații mai clare pentru cei care operează cu valorile transmisivității în modelele numerice de simulare a dinamicii acviferelor.

COMENTARIU

Optimizarea rețelei de cercetare este operațiunea care utilizează toate rezultatele modelării geostatistice ale structurilor spațiale.

După parcurgerea tuturor etapelor de prelucrare, plecând de la premiza că toate au fost corect realizate, ne putem manifesta, pe baza unei fundamentări cantitative complete, acordul sau dezacordul în legătură cu rezultatul.

Elementele principale care trebuie luate în considerare sunt:

- *reprezentativitatea datelor pentru distribuția spațială a variabilei studiate;*
- *gradul de certitudine al estimărilor reflectat în valorile erorilor de estimare;*
- *posibilitatea îmbunătățirii preciziei estimărilor pe baza optimizării rețelei de cercetare.*

Și iată-ne cu tabloul structurii în față.

*Dacă nu ne place cum arată o putem lua de la capăt dar nu oricum. Există un control al imaginației prin intermediul datelor utilizate. Orice estimare este caracterizată de un anumit grad de adecvare al **modelului structural** (modelul de variogramă) și de **precizia de estimare** (abaterea standard de estimare).*

Dacă rezultatul obținut nu este în concordanță cu așteptările, în condițiile unor date considerate reprezentative, este cazul să suspectăm de falsitate ipotezele pe care vrem să le verificăm.

Au apus vremurile în care imaginația și flerul geologului rezolvau mare parte din problemele cercetării geologice.

Datele măsurate corect și metodele cantitative aplicate cu rigoare sunt singurele care pot să confirme sau infirme ipoteze pe baza cărora se proiectează exploatarea unor zăcăminte de petrol, se amplasează captarea necesară alimentării cu apă a unei localități sau se decide dacă un zăcământ de aur este rentabil sau nu.

*Nu încercați să construiți o hartă geologică fără o metodologie cantitativă bine fundamentată! **Nu pictați** hărți sau secțiuni geologice cu pensula, chiar dacă pensula o manevești cu “mouse-ul”!*

O hartă este adevărată nu pentru că o desenați utilizând programul pentru desenare COREL DRAW, ci pentru că o calculați utilizând programul de estimare geostatistică Geo-EAS!

BIBLIOGRAFIE

Andrews, D.J. & Hanks, T.C., *Scarp degraded by linear diffusion : inverse solution for age*, J.Geophys.Res.90, 10193-208, 1985.

Bailey, N.T.J., *The elements of stochastic processes with applications to the natural sciences*, John Wiley & Sons, Inc., New York, 1964.

Bergé, P., Poneau, Y. & Vidal, C., *Order within chaos*, John Wiley and sons, New York, 1986.

Bomboe, P., *Geologie matematică (vol. I, Analiza statistică a datelor geologice)*, Editura Universității din București, 1979.

Brown, S.R., *A note on the description of surface roughness using fractal dimension*, Geophys. Res. Lett. 14, 1095-8, 1987.

Cennini, C., *Tratatul de pictură*, Ed.Meridiane, 1977.

Chauvet, P., *Aide memoire de Geostatistique Lineare*, Fascicule 2, Cahiers de Geostatistique, Centre de Geostatistique, Ecole de Mines de Paris, 1991.

Cheaney, R.F., *Statistical methods in geology*, George Allen & Unwin (publishers) Ltd, London, 1983.

Clarke, G.P.Y. and Dane, J.H., *A simplified theory of point kriging and its extension to cokriging and sampling optimization*, Bulletin 609, Alabama Agricultural Experiment Station, Auburn University, Alabama, february 1991.

Craiu, V., Enache, R., Bâscă, O., *Teste de concordanța cu programe în Fortran*, Editura științifică și enciclopedică, București, 1986.

Daccord, G. & Lenormand, R., *Fractal patterns from chemical dissolution*, Nature 325, 41-3, 1987.

David, M., *Handbook of applied advanced geostatistical ore reserve estimation*, Elsevir, Amsterdam, 1988.

David, M., *Geostatistical ore reserve estimation*, Elsevier, Amsterdam, 1977.

Davis, J. C., and McCullagh, M. J., *Display of analysis data*, Wiley, New York, 1975.

Delfiner, P., Matheron, G., *Les fonction Aleatoires Intrinseques d'ordre k*, Les Cahiers du Centre de Morphologie Mathematique de Fontainebleau, Ecole de Mines de Paris, 1980.

Delhomme, J.P., *Les variables regionalisees dans les sciences de l'eau*, B.R.G.M., Deuxieme serie, no4, Section III, Hydrogeologie-geologie de l'ingeneur, Paris, 1978.

Deutsch, C.V., Journel, A.G., *GSLIB: Geostatistical Software Library*, New York, Oxford University Press, 1992.

Deverle, P., H., *Mineral resources appraisal*, Calderon Press, Oxford, 1984.

Dick O., *Fractalvision : Put fractals to work*, București, Teora, 1995.

Dubuc, B., Quiniou, J.F., Roques-Carmes, C., Tricot, C. & Zucker, S.W., *Evaluating the fractal dimension of profiles*, Phys.Rev. A39, 1500-2, 1989.

Fabbri, A.G., *Image processing of geological data*, New York, Van Nostrand reinhold Company, 1984.

Fabbri, A.G., and Kasvand, T., *Image processing for detection of two-dimensional markovian prpperties as functions of distances from crystal profiles*, in Proc. 3rd European symposium dor stereology , Ljubliana, Yugoslavia, June 22-26, 1981, Stereologia Iugoslavica, v. 3, (suppl. 1),

Fouquet, Ch.De, *Simulation conditionnelle de fonctions aleatoires: cas gaussien stationnaire et schema lineaire*, Centre de Geostatistique, Ecole des mines de Paris, 1993.

Guillaume, A., *Analyse des variables regionalise*, Doin Editeur, Paris, 1977.

Hirata, T., Satoh, T. & Ito, K., *Fractal structure of spatial distribution of microfracturing in rock*, Geophys. J. Roy. Astron. Soc. 90, 369-74, 1987.

Houlding, S.W., *Practical Geostatistics, Modeling and Spatial Analysis*, Springer-Verlag Berlin Heidelberg, 2000

Isaaks, E.H., Srivastava, M.R., *An introduction to Applied Geostatistics*, New York, Oxford University Press, 1989.

Journel, A.G., Huijbregts, Ch.J., *Mining Geostatistics*, Academic Press, London, 1978.

Journel, A.G., *Exploitation des mines. Guide pratique de geostatistique*, Ecole des mines d'Ales, 1975.

Kasvand, T., Fabbri, A.G. and Nel, L.D., *Digitization and processing of large regional geological maps*, Nat. Res. Council Can., Elec. Eng. Division, Report, ERB-938, 1981.

Kecs, W., *Complemente de matematică cu aplicații în tehnică*, Editura tehnică, București, 1989.

Kruhl, J.H., *Fractals and dynamic systems in geoscience*, Springer-Verlag, Berlin Heidelberg New York., 1994.

Laffite, P., *Traité d'informatique géologique*, Masson et Cie Editeurs, Paris.

Marsily, G.De, *Quantitative Hydrogeology*, New York, London, Academic Press, INC, 1986.

Matheron, G., *Traite de Geostatistique Appliquee*, (tome I), Technip, Paris, 1976.

Matheron, G., *Traite de Geostatistique Appliquee*, (tome II), Technip, Paris, 1963.

Matheron, G., *La theorie des variables régionalisées, et ses applications*, Les Cahiers du Centre de Morphologie Mathematique de Fontainebleau, Fascicule 5, Ecole de Mines de Paris, 1970.

Matheron, G., *Le choix des modèles en géostatistique*, in *Advanced Geostatistics for mining industry.*, Guaracio et al., Reidel, 1976.

Matheron, G., *Estimer et choisir*, Les Cahiers du Centre de Morphologie Mathematique de Fontainebleau, Fascicule 7, Ecole de mines de Paris, 1978.

McCall, J., and Marker, B. (editors), *Earth science mapping*, Graham & Trotman, London, 1989.

Mihoc, G.m Bergthaller, C., Urseanu, V., *Procese stocastice*, Editura științifică și enciclopedică, București, 1978.

Mont, O'L., Lippert, R. H., Spitz, O.T., *Fortran IV and map program for computation and plotting of trend surfaces degrees 1 through 6*, Michigan, 1979.

Murgu, M., *Analiza rețelelor de explorare și valorificarea optimă a zăcămintelor minerale*, Tipografia Univ. București, 1979.

Onicescu, O., Ștefănescu, V., *Elemente de statistică informațională cu aplicații*, Editura tehnică, București, 1979.

Preston, F.W., and Davis, J.C., *Sedimentary porous materials as a realization of stochastic processes*, in *Random Processes in Geology*, D.R. Merriam, ed., Springer-Verlag, New-York, 1976.

Rivoirard, J., *Introduction au krigeage disjonctif et a la geostatistique non lineaire*, Centre de Geostatistique, Ecole des mines de Paris, 1991.

Rosenfeld, A., & Kak, A.C., *Digital picture processing*, Academic press, New york, 1976.

Rousseau, J.J., *Scrieri despre artă*, B.P.T., București, 1981.

Schwarzacher, W., *Sedimentation models and quantitative stratigraphy*, Elsevier scientific publishing company, Amsterdam, 1975.

Scrădeanu, D., Mihnea, G., *L'etude de variations spatiales de grandeurs hydrogeologique a l'aide du krigeage*, Analele Univ.Bucuresti, 1987.

Scrădeanu, D., *Optimizarea metodelor de explorare a zăcămintelor de lignit*, Teză de doctorat, Univ.Buc, 1993.

Scrădeanu, D., *Informatică geologică*, Editura Univ.Bucuresti, 1995.

Scrădeanu, D., *Modele geostatistice în Hidrogeologie*, vol.I, Editura didactică și Pedagogică, R.A.-București, 1996.

Shakeel, A., *Estimation des transmissivites des aquifers par methodes geostatistique mulrivariables et resolution indirecte du probleme inverse*, These presentee a l'Ecole Nationale Superieure des Mines de Paris, 1987.

Silași, I., *Geostatistică aplicată în cercetarea zăcămintelor și evaluarea rezervelor*, Multiplicat în atelierile C.P.P.G. al M.M.P.G.,București, 1975.

Srivastava, G. S., *Optical processing of structural contour maps*, J. Math. Geol. 9, 1975.

Strang, G., *Linear algebra and its applications*, Academic Press, New York, 1980.

Teodorescu, D., *Modele stohastice optimizate*, Editura Academiei R.S.R, București, 1982.

Trescott, P.C. et al., *Finite-difference model for aquifer simulation in two dimensions with results of nuerical experiments*, Geological Survey, Washington, 1976.

Turcotte D.L., *Fractals chaos in geology and geophysics*, Cambridge University Press, 1992.

Wackernagel, H., *Cours de geostatistique multivariable*, Centre de Geostatistique, Ecole des mines de Paris, 1993.

Wiener, U., Isaic-Maniu, A., Vodă, V., *Aplicații ale rețelelor probabiliste în tehnică*, Editura tehnică, București, 1983.

Tatarkiewicz,W., *Istoria esteticii*, Editura meridiene, București, 1978.

Zorilescu, D., *Proгноza resurselor de materii prime minerale*, Editura tehnică, București, 1975.

Zorilescu, D., *Modele operaționale ale problemelor miniere*, Editura tehnică. București, 1981.

Zorilescu, D., *Introducere în geostatistica informațională*, Editura Academiei, București, 1990.

4.GHID DE UTILIZARE A PACHETULUI DE PROGRAME Geo-EAS

Pachetul de programe Geo-EAS (**Geostatistical Environmental Assessment Software**) conține 13 subprograme utile pentru analiza distribuției spațiale a *caracteristicilor geologice*, probate într-o rețea de explorare cu geometrie neregulată/regulată.

Cele 13 subprograme sunt:

- **Dataprep** - pentru operații cu fișierele care conțin datele;
- **Trans** - pentru transformarea fișierelor de date;
- **Stat1** - pentru analiză statistică univariată;
- **Scatter** - pentru realizarea hărților punctuale și a dreptelor de regresie;
- **Prevar** - pentru pregătirea calculului variogramei experimentale;
- **Vario** - pentru calculul variogramei experimentale și modelarea ei;
- **Xvalid** - pentru validarea modelului variogramei experimentale;
- **Krige** - pentru estimarea punctuală și zonală prin kriging;
- **Conrec** - pentru realizarea hărților conturale;
- **Xygraph** - pentru diagrame rectangulare;
- **Postplot** - pentru hărți punctuale;
- **Hplot** - pentru transferul fișierelor care conțin grafice;
- **View** – pentru vizualizarea graficelor din fișiere de tip “metacode.”

Subprogramele pot fi activate dintr-un menu principal (*Fig.4.1*) sau din fișierele executabile. Acest menu se activează din fișierul **geoeas.exe**. Pentru a utiliza unul din subprograme se procedează în modul următor:

- se poziționează cursorul pe numele subprogramului respectiv (folosind săgețile direcționale : ←, ↑, →, ↓);
- în partea de jos a ecranului se poate citi o scurtă descriere a programului;
- se tastează <Enter>.

GEO -EAS (1.0)		
Geostatistical Environmental Assessment Software		
Dataprep	Prevar	Conrec
Trans	Vario	Xygraph
Stat1	Xvalid	Postplot
Scatter	Krige	HPplot
		View
		Quit
Use arrow to move cursor, <enter> to select program		
Data preparation utilities		

Fig.4.1. Menu principal

4.1. FIȘIERELE DE DATE

Fișierele care conțin datele inițiale pentru subprogramele Geo-EAS sunt de tip ASCII și au o structură specifică.

4.1.1. STRUCTURA FIȘIERELOR

Structura fișierului de tip Geo-EAS este compusă din două părți (**Tabelul 4.1**):

- *identificatorul variabilelor*;
- *corpul valorilor*.

Identificatorul variabilelor permite programului Geo-EAS să coreleze numele variabilelor cu valorile lor numerice și este compus din mai multe linii:

- *Linia 1*, de maxim **256 caractere** care de regulă conține un comentariu ce precizează proveniența datelor din fișier.
- *Linia 2*, pe care este precizat *numărul variabilelor (NVAR)* din fișier. Fișierele acceptate de versiunea care circulă pe Internet pot conține maximum **48 de variabile** (coloane) și **10000 valori pentru fiecare variabilă**.
- *Liniile 3 până la NVAR+2* pe care se precizează: *Numele fiecărei variabile și unitatea de măsură* corespunzătoare. Se folosește câte o linie pentru fiecare variabilă. Numele variabilei poate avea maxim **10 caractere** care se scriu pe primele 10 coloane iar de pe a **11-a coloană** până al a **20-a** se trece unitatea de măsură (opțional). Numele variabilelor și unitatea de măsură sunt folosite ca text pentru reprezentările grafice.

Corpul valorilor se scrie începând din linia *NVAR +3*. Datele se scriu pe coloane, separate prin *virgulă* sau printr-un spațiu (**Nu cu TAB!!!!**). Valorile variabilelor trebuie să fie numerice. Dacă, dintr-un punct de observație valoarea unei variabile lipsește, pe poziția ei din fișier se scrie valoarea *1.E31* pe care programele Geo-EAS nu o iau în calcul.

Tabelul 4.1. Fișier de date Geo-EAS

H5.dat - Acviferul din baza stratului V de carbune					
4					" identificatorul variabilelor"
X	m				
Y	m				
Z_NH	m				
T	m ² /zi				
54305.65	70960.68	140.93	16.815		"corpul valorilor"
53389.32	72456.37	151.9	1.428		
54534.84	74103.48	168.12	1.75		
50866.36	73018.48	164.87	0.0288		
50162.36	70618.48	144.87	1.E+31		
54004.84	71103.48	136.12	1.41		
...		

4.1.2. CONVENȚII PENTRU EXTENSIA FIȘIERELOR

Deși programele din Geo-EAS nu au restricții relativ la extensia fișierelor este recomandabilă utilizarea unor extensii convenționale pentru diferitele categorii de fișiere create pe parcursul prelucrării.

Principalele categorii de extensii ale fișierelor pe care le utilizează subprogramele Geo-EAS sunt:

TXT - fișier în format ASCII;

DAT - fișier de date tip Geo-EAS

PCF - fișierul creat de *PREVAR* pentru calculul variogramei experimentale;

GRD - fișier creat cu subprogramul *KRIGE*;

CPF - fișier cu parametri pentru subprogramul *CONREC*;

KPF - fișier cu parametri pentru subprogramul *KRIGE*;

XPF - fișier cu parametri pentru subprogramul *XYGRAPH*;

POL - fișier cu coordonatele poligonului contural pentru subprogramul *KRIGE*;

MET - fișier grafic (metacode) creat de *CONREC*, *POSTPLOT*, *XYGRAPH*;

PLT - fișier cu instrucțiuni de imprimare creat de *HPLOT*.

4.2. ECRANE INTERACTIVE

Toate subprogramele Geo-EAS au caracteristici interactive similare cuprinse într-un cadru grafic unitar (**Fig.4.2**) ce poate fi separat în patru categorii: *cadrul ecranului (A)*, *linia pentru mesaje (B)*, *linia cu opțiunile de prelucrare (C)*, *grupurile de parametri (D)*.

A.Cadrul ecranului este un dreptunghi în care sunt cuprinse câmpurile pentru introducerea datelor și afișarea rezultatelor. Acest cadru este subîmpărțit pentru structurarea comenzilor de lucru.

B.Linia pentru mesaje apare la partea inferioară a cadrului ecranului și conține mesaje în legătură cu luarea unei decizii privind continuarea prelucrării.

Exemple:

1. La citirea unui fișier neadecvat apare mesajul “eroare la citirea identificatorului de variabile” (engl. “error reading header”). Pentru continuarea prelucrării se iese din program și se corectează fișierul de date.

2. În cazul în care nu se găsește fișierul de date în directorul indicat apare mesajul “fișierul nu este găsit” (engl. “file not found”).

C. Linia cu opțiunile de prelucrare (engl. “menu line”) este poziționată exact sub cadrul ecranului. Când se alege o anumită opțiune (prin plasarea cursorului pe numele opțiunii respective) sub menu apare o scurtă descriere a opțiunii respective. Pe ecranul principal se va găsi și o descriere mai detaliată a opțiunii respective. Pentru a selecta o opțiune se tastează litera majusculă din numele opțiunii sau se poziționează cursorul pe numele ei și se tastează <enter>.

Exemplu:

În programul *STAT 1* pentru a alege opțiunea de introducere a numelui fișierului de date se tastează litera *D*, prima din numele opțiunii *Data*.

D. Grupurile de parametri

Fiecare grup de parametri se accesează prin intermediul menu-ului. Grupurile de parametri sunt legate funcțional. Când un grup de parametri este accesat, în interiorul chenarului care cuprinde parametrii apare un câmp activ în dreptul unuia dintre ei. Acest câmp activ se mută de la un parametru la altul cu ajutorul săgeților. După ce s-au introdus toți parametrii din grup, cursorul se poziționează automat în menu-ul principal.

STAT1 (1.0)	
A programe to compute univariate descriptive statistics	
File Prefix: C:\GeoEAS\Data\ (A)	
File Data File Name: H5.dat	# Variables : 4 # Data records : 131 # Missing Data: 0
Variable: Variable : Z_NH Weight : None (D) Log Option : Off	Execute Use this options to compute and display the statistics.
Limits: Minimum : .000 Maximum : 158.00	A menu of additional graphs and options will be displayed.
(B)	
Prefix Data Variable Limits Compute basic descriptive stats	Execute Batch Statistics Quit (C)

Fig.4.2.Exemplu de ecran interactiv

Introducerea parametrilor se face în trei categorii de câmpuri: *câmpuri alfanumerice*, *câmpuri numerice* și *prompt-uri yes/no*.

Câmpurile alfanumerice, (*Prefix* sau *Data* din *STAT1*) acceptă caractere alfabetice sau numerice (identificatorul directorului în care se află fișierul de date, identificatorul fișierului cu date).

Câmpurile numerice se alimentează numai cu caractere numerice. Dacă se introduc valori alfanumerice programul atenționează printr-un sunet specific. Un exemplu de câmp numeric este cel de la opțiunea *Limits* din *STAT1*. Se pot introduce caracterele numerice de la <0> la <9> și <.>.

Câmpurile alternative sunt câmpuri specifice care conțin o listă de 2 sau mai multe variante. Aceste variante se pot alege folosind tasta "space". Când în câmpul alternativ apare valoarea dorită aceasta se transmite programului prin <enter>.

Exemple:

1.În câmpul "Variable", la utilizarea tastei "space" vor apare succesiv toate variabilele din identificatorul de variabile al fișierului de date. Când variabila dorită apare în câmp se tastează <enter>.

2.În câmpul "Log" apar două opțiuni "On" și "Off". În varianta "On" se vor logaritma valorile din fișierul de date pentru calculele ulterioare.

Prompt-uri yes/no care sunt utilizate pentru informații adiționale. Acestea sunt

folosite pentru a obține informații care nu pot fi afișate permanent pe ecran. Aceste opțiuni vor apărea temporar pe linia de mesaje și au forma “Întrebare.....<Y/N>?”. Pentru "Da" se apasă “Y” iar pentru "Nu": “N”.

Exemplu: “Do you really want to quit <Y/N>?”.

Unele opțiuni din menu sunt urmate de câmpuri pentru informații adiționale care apar pe linia de mesaje și pot fi *alfanumerice*, *numerice* sau *alternative*.

4.3.STRUCTURA SUBPROGRAMELOR

Subprogramele din sistemul Geo-EAS cer date din fișiere tipice și opțiuni prin interacțiune, de la tastatură. Aceste subprograme sunt aranjate într-o anumită ierarhie sau în grupuri cu funcționalitate asemănătoare. Fiecare grup sau valoare a unui parametru dintr-un subprogram este accesată printr-un menu de opțiuni. Unele opțiuni conduc la alt menu în timp ce altele la un dialog în care urmează să se introducă parametri. Unele opțiuni conduc la cerere de date pentru subprogram iar altele la rezultate numerice sau grafice. Identificarea ierarhiei și opțiunilor subprogramelor se face prin traversarea menu-ului și citirea mesajelor explicative asociate.

4.4.OPȚIUNI COMUNE

Opțiunile comune programelor din pachetul Geo-EAS sunt: ***Prefix***, ***Data***, ***Variable(s)***, ***Execute***, ***Read Parameters***, ***Save Parameters***, ***Quit***.

Prefix este opțiunea prin care se introduc drive-ul și identificatorul directorului de lucru (de unde se iau fișierele de date și unde se vor salva diferitele fișiere cu rezultatele prelucrării).

Data este opțiunea prin care introducem identificatorul fișierului de date care urmează să fie folosit pentru prelucrarea cu subprogramele din Geo-EAS. Orice eroare care apare în timpul citirii fișierului este semnalată printr-un mesaj explicit. Dacă nu apare nici un mesaj de eroare numele variabilelor citite din fișierul de date sunt înmagazinate în câmpurile alternative pentru a putea fi selectate cu opțiunea ***Variables***.

Variable este o opțiune comună pentru toate subprogramele Geo-EAS. Ea permite alegerea unei/unor variabile care urmează să fie prelucrate. Unele subprograme utilizează numai o variabilă (ex: *Stat1*) iar altele chiar trei (ex.: *Conrec*). În mod normal această opțiune conduce la un câmp alternativ care conține numele variabilei, variabilă care poate fi selectată cu ajutorul tastei “*space*”.

Execute este opțiunea comună tuturor subprogramelor din Geo-EAS. Ea este folosită pentru a iniția prelucrarea datelor.

Opțiunile ***Read Parameters*** și ***Save Parameters***, sunt comune subprogramelor

care utilizează fișiere cu parametri. Aceste fișiere conțin valori pentru toate opțiunile de prelucrare. Parametrii salvați pot fi folosiți ulterior cu opțiunea *Read parameters*. Alegerea acestor opțiuni conduce la un câmp în care trebuie introduse identificatorul și directorul fișierelor cu parametri. **Fișierele cu date și cele cu parametrii aleși trebuie să fie în același director.**

Opțiunea **Quit** este folosită pentru a ieși din program. Această opțiune este valabilă pentru a coborî un nivel în structura arborescentă. Poate fi folosită și tasta "Q" pentru a executa aceeași operație.

4.5 GRAFICĂ Geo-EAS

Subprogramele *Stat1*, *Vario*, *Xvalid* și *Krige* realizează graficele direct pe monitor. Aceste grafice pot fi imprimate imediat folosind tasta "P" dar nu pot fi salvate într-un fișier. Dacă imprimanta căreia i se transmite graficul nu este "on line" programul se "blochează" și este necesară reinițializarea calculatorului.

Subprogramele *Postplot*, *Xygraph* și *Conrec* creează grafice pe monitor după ce au creat un fișier "metacode". Fișierul "metacode" poate fi folosit ulterior pentru imprimare sau vizualizare, el conținând informații referitoare la un anumit tip de imprimantă sau monitor. Subprogramul *View* traduce fișiere metacode pentru monitoare EGA, CGA și Hercules. Programul *Hpplot* transformă fișierele metacode pentru imprimante Hewlett Packard. Calitatea graficelor executate cu ajutorul fișierelor metacode este mai bună decât cea a graficelor produse cu programe non-metacode.

4.6.ETAPELE UNUI STUDIU GEOSTATISTIC ȘI PROGRAMUL Geo-EAS

Sucesiunea utilizării subprogramelor pachetului Geo-EAS pentru parcurgerea etapelor de prelucrare ale unei *estimări geostatistice* este următoarea:

- 1.Post Plot** permite realizarea hărților punctuale necesare descrierii distribuției spațiale a datelor prelucrate.
- 2.Stat1** facilitează descrierea variabilității globale a caracteristicilor studiate (tipul de repartiție a frecvențelor valorilor prelucrate, parametrii statistici elementari etc.).
- 3.Prevar** realizează fișierul pentru calculul variogramei experimentale. Acest fișier este salvat în format binar și în mod uzual are extensia "pcf" (acronimul pentru : *pair comparison file*).
- 4.Vario** este subprogramul cu care se calculează și modelează variograma experimentală. Modelul variogramei experimentale obținut controlează calitatea rezultatelor interpolării realizată prin kriging.

ATENȚIE!!!

ALEGEREA MODELULUI DE VARIOGRAMĂ ESTE CEA MAI IMPORTANTĂ DECIZIE A ESTIMĂRII STRUCTURILOR SPAȚIALE !

Pentru calculul variogramei experimentale este necesar fișierul creat în Prevar (fișierul cu extensia "pcf"). Numărul de variabile, de probe și de perechi de puncte utilizabile este limitat în funcție de versiunea programului.

5.Xvalid permite validarea modelului ales pentru variograma experimentală. Acest subprogram compară valorile măsurate în punctele de probare cu cele estimate prin kriging, pe baza modelului de variogramă. Valabilitatea modelului se evaluează pe baza diferențelor dintre valorile măsurate și cele calculate.

6.Krige este subprogramul care pe baza *modelului de variogramă* ales și al *valorilor măsurate* în punctele de observație realizează estimarea valorii caracteristicii în orice punct din zona cercetată.

7. Postplot, Xygraf, Conrec, View și Hplot sunt subprogramele pachetului Geo-EAS care permit realizarea reprezentării grafice a rezultatelor estimării structurilor spațiale.

4.7. PRINCIPALELE SUBPROGRAME Geo-EAS

Principalele subprograme din pachetul Geo-EAS care sunt indispensabile estimărilor geostatistice sunt *Stat1*, *Vario* și *Krige*.

4.7.1. STAT1

Subprogramul *Stat1* utilizat pentru studiul distribuției valorilor variabilelor prelucrate are structura prezentată în **Fig.4.3**.

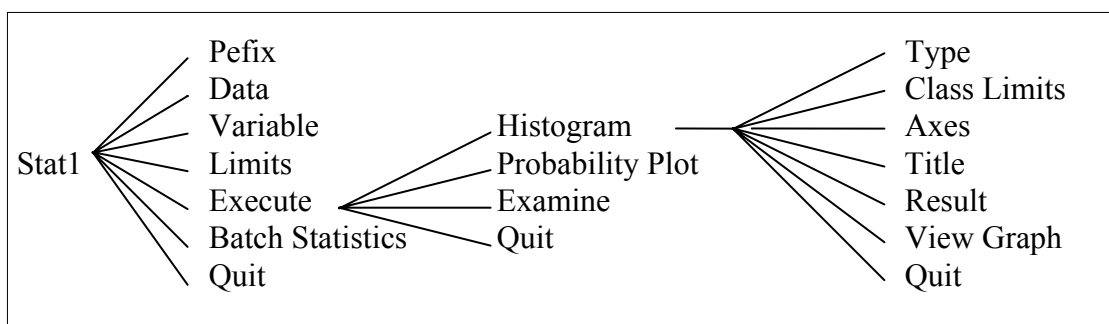


Fig. 4.3. Structura subprogramului STAT1

Posibilitățile de prelucrare ale opțiunilor specifice subprogramului sunt:

- **Variable** permite selecționarea variabilelor și a modului de prelucrare a valorilor (valori ponderate sau neponderate, logaritmă sau nelogaritmă,);
- **Limits** permite selecționarea valorilor cuprinse între o valoare minimă și una maximă definite de utilizator.
- **Execute** permite construirea graficelor *histogramei*, *diagramei de probabilitate*, precum și examinarea valorilor numerice pe baza cărora s-au realizat acestea.
- **Histogram** permite realizarea unor opțiuni de reprezentare și analiză prin:
 - **Type** - alegerea tipului de histogramă cu frecvențe absolute/relative;
 - **Class Limits** - alegerea mărimii intervalului de grupare;
 - **Axes** - modificarea textului asociat axelor;
 - **Title** - modificarea titlului asociat cu graficul histogramei;
 - **Result** - afișarea frecvențelor pe baza cărora este construită histograma.
- **Batch Statistics** este un fișier care conține toate rezultatele prelucrării statistice univariate realizate cu *Stat1*.

4.7.2.VARIO

Parcurgerea structurii ierarhizate a programului *Vario* (Fig.4.4) permite calculul și modelarea variogramei experimentale.

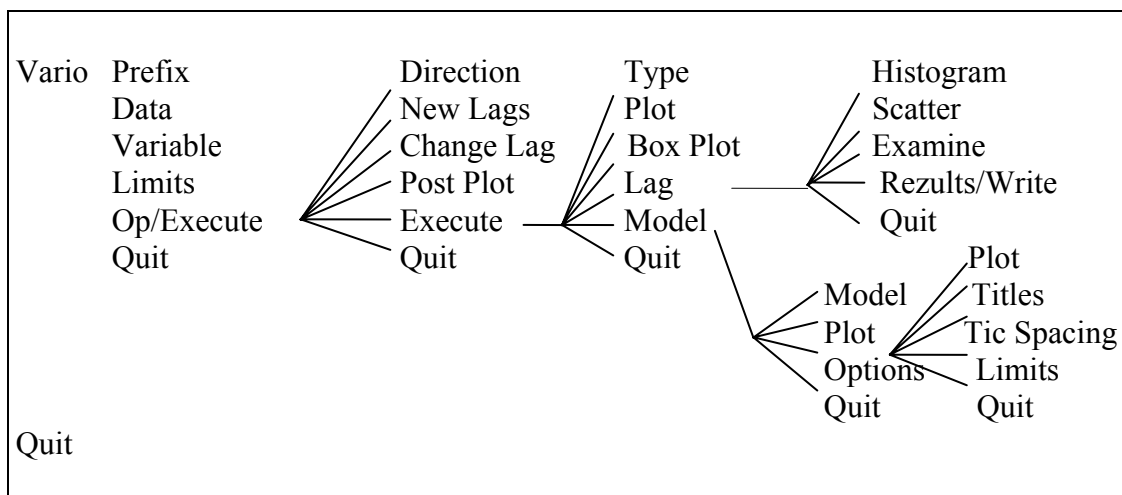


Fig. 4.4. Structura subprogramului *Vario*.

După parcurgerea opțiunilor comune (*Prefix*, *Data*, *Variable*, *Limits*), *Option/Execute* deschide sesiunea de calcul a variogramei experimentale cu:

- **Direction** pentru stabilirea direcției de calcul a variogramei, toleranței de direcție și zonei de căutare. Orientarea direcției de calcul se exprimă prin unghiul dintre aceasta și direcția vest-est (în grade sexagesimale).
- **Tolerance** permite introducerea toleranței de direcție, cu valori cuprinse între 0° și 90° .

• **Max Band-width** permite precizarea lățimii zonei asociate direcției din care sunt selecționate perechi de puncte pentru calculul variogramei (Fig.4.5.)

• **New Lags** permite alegerea claselor de distanțe pentru calculul variogramei prin intermediul unor opțiuni subordonate:

- **Minimum** - valoarea minimă a distanței dintre perechile de puncte;
- **Maximum** - valoarea maximă a distanței dintre perechile de puncte;

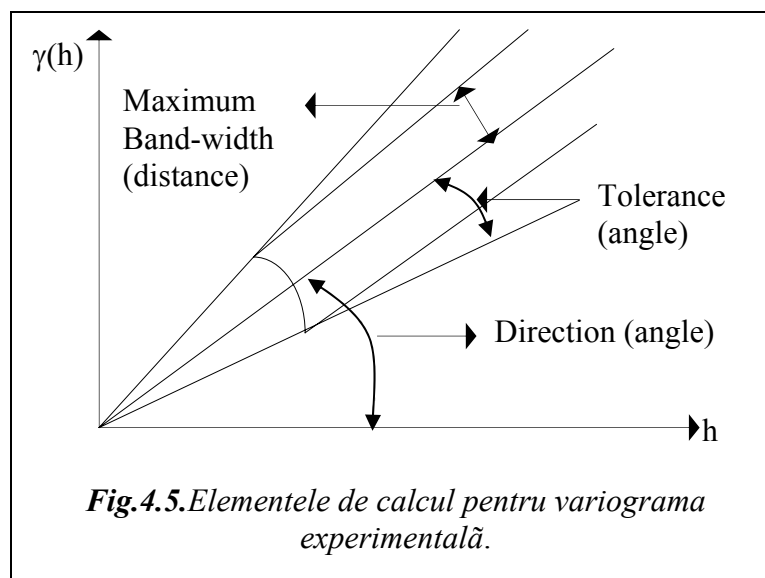


Fig.4.5. Elementele de calcul pentru variograma experimentală.

Increment - pasul de creștere al distanței dintre perechile de puncte.

După precizarea parametrilor de calcul se calculează valorile variogramei experimentale prin selecționarea opțiunii **Execute**, opțiune în care sunt integrate o serie de alte opțiuni pentru reprezentarea grafică a rezultatelor calculului:

- **Type** permite alegerea tipului de funcție de distanță care se calculează: *covarianța inversă (InvConv)*, *variograma (Vario)*, *variograma relativă (Relative)*, *madograma (Madogram)*.

- **Plot** realizează reprezentarea grafică a funcției de distanță selecționată. Pentru continuarea prelucrării, după afișarea graficului se tastează *Q* și se revine la rezultatele calculului funcției de distanță.

- **BoxPlot** realizează reprezentarea grafică a distribuției valorilor funcției de distanță calculate pentru fiecare clasă de distanță.

- **Lag Results** permite selecționarea unei *clase de distanță (Lag)* și analiza distribuției lor printr-o serie de alte opțiuni subordonate:

- **Histogram** realizează histograma valorilor;
- **Scatter** realizează diagrama de continuitate;
- **Examine** afișează pentru fiecare clasă de distanță: perechile de puncte utilizate (*Pair*), valorile măsurate în cele două puncte (*1st value, 2nd value*), distanța dintre cele două puncte (*Distance*), azimutul direcției segmentului care unește cele două puncte (*Direction*), pătratul diferenței dintre valorile măsurate în cele două puncte (*Diference^2*). Renunțarea la tabelul cu valorile afișate de **Examine** se face cu tasta *Q*.
- **Write** scrie rezultatele obținute într-un fișier de tip ASCII al cărui nume este stabilit de utilizatorul programului.
- **Quit** permite revenirea la tabelul cu valorile funcției de distanță.

- **Model** este opțiunea care permite stabilirea interactivă a modelului printr-o serie de alte opțiuni subordonate:

- **Model** permite introducerea tipului modelului și a parametrilor acestuia:
 - efectul de pepită (*Nugget*);
 - tipul (*Type*) care poate fi ales utilizând tasta "space" dintr-un repertoriu de patru modele: sferic (*Spherical*), gaussian (*Gaussian*), exponențial (*Exponent*), linear (*Linear*).

- **Plot** suprapune peste graficul variogramei experimentale graficul modelului ales. Opțiunile *Model* și *Plot* se utilizează succesiv până când graficul modelului realizează cea mai bună calare a variogramei experimentale. Se pot suprapune maximum patru modele analitice pentru calarea modelului de variogramă.

- **Options** oferă posibilitatea modificării reprezentărilor grafice realizate de subprogram cu parametrii implicați prin:

- **Titles** care permite modificarea textului atașat reprezentării grafice a variogramei experimentale și modelului (titlul, denumirea variabilelor de pe axele de coordonate);
- **Tic Spacing** utilizat pentru precizarea scării grafice a axelor de coordonate;
- **Limits** este utilizat pentru stabilirea amplitudinii variabilelor de pe cele două axe de coordonate (*X Axis Minimum, X Axis Maximum, Y Axis Minimum, Y Axis Maximum*).
- **Quit** este opțiunea de ieșire din toate nivelurile de prelucrare.

4.7.3. XVALID

Subprogramul *Xvalid* validează modelul de variogramă ales prin compararea valorilor calculate prin kriging (pe baza modelului de variogramă) cu cele măsurate.

Structura subprogramului este organizată pe trei niveluri (**Fig.4.6**) cu următoarele opțiuni:

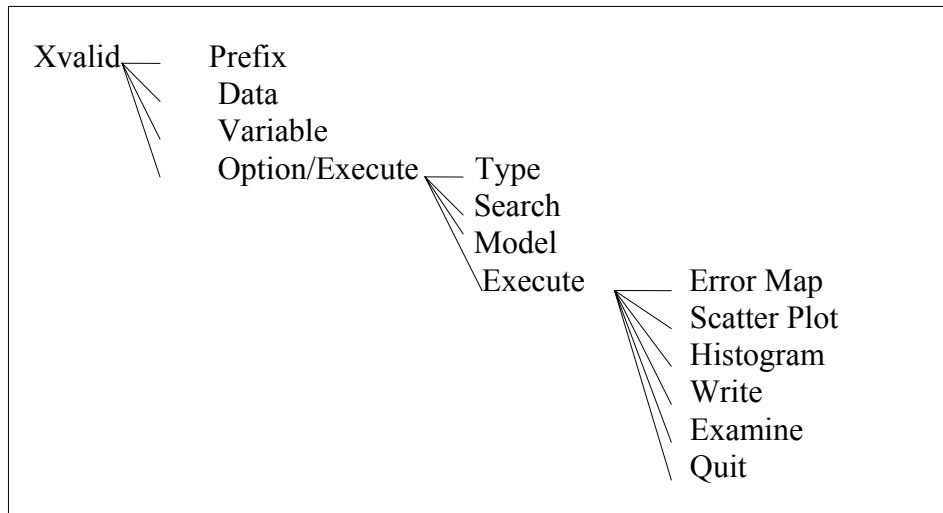


Fig.4.6. Structura subprogramului *Xvalid*

- **Prefix, Data și Variable** care permit precizarea directorului de lucru, identificatorul fișierului de date și variabila care se prelucrează;
- **Option/Execute** permite accesul la opțiunile necesare precizării parametrilor necesari validării modelului de variogramă;
 - **Type** pentru alegerea tipului de kriging utilizat: simplu (*Simple*) sau ordinar (*Ordinary*);
 - **Search** delimitează mărimea, forma și structura zonei de calcul pentru fiecare punct verificat prin: raza mare a zonei de căutare (*Major Radius*), raza minimă de căutare (*Minor Radius*), unghiul dintre raza mare și axa OX (*Angle*), numărul de sectoare al elipsei (*Sectors*). Pentru creșterea eficienței calculului se pot stabili numărul de sectoare admise, fără puncte (*Empty sectors*), numărul maxim de puncte din fiecare sector (*Max in Sector*), precum și cel minim (*Min. to use*).
 - **Model** permite introducerea parametrilor modelului variogramei care urmează să fie validat: efectul de pepită (*Nugget*), tipul modelului (*Type*), valoarea maximă a variogramei (*Sill*), razele de influență (*Major Range, Minor Range*), unghiul dintre raza mare și direcția OX (*Ellipse Angle*).
 - **Execute** este opțiunea care declanșează operațiunile de estimare ale valorii variabilei prelucrate în fiecare punct de observație pe baza modelului de variogramă ales. După realizarea calculului pentru toate punctele de observație se trece automat într-un ecran care prezintă rezultatele unei analize statistice globale (*media, dispersia, abaterea standard și quartilele* diferențelor dintre valorile estimate și cele măsurate) și opțiuni de analiză detaliată a modelului de

variogramă testat:

- **Error Map** realizează o hartă a distribuției erorilor în zona studiată. Erorile de supraestimare sunt reprezentate prin "+"-uri iar cele de subestimare prin "x"-uri. Dimensiunile semnelor sunt proporționale cu valoarea absolută a erorilor.
- **Scatter Plot** realizează două diagrame binare rectangulare:
 - valorile estimate (*Kriged Estimate*) raportate la valorile măsurate (*Variable*);
 - valorile estimate (*Kriged Estimate*) raportate la diferența dintre valorile estimate și cele măsurate (*Difference: Estimate-Variable*).
- **Histogram** permite reprezentarea distribuției frecvențelor erorilor de estimare ($Error=Kriged\ Estimate-Variable$) și a erorilor reduse ($(Kriged\ Estimate-Variable)/Ksdev$).
- **Write** permite scrierea rezultatelor validării într-un fișier de tip ASCII.
- **Examine** permite examinarea rezultatelor validării pe monitor.
- **Quit** este pentru ieșirea din programul de validare.

4.7.4.KRIGE

Subprogramul *Krige* realizează pe suprafața cercetată o rețea rectangulară în ale cărei noduri sunt calculate *valorile estimate* prin kriging “ordinar” sau “simplu” și *abaterile standard*. Rețeaua implicită este de 10 x 10 noduri. Structura subprogramului este reprezentată în **Fig. 4.7**.

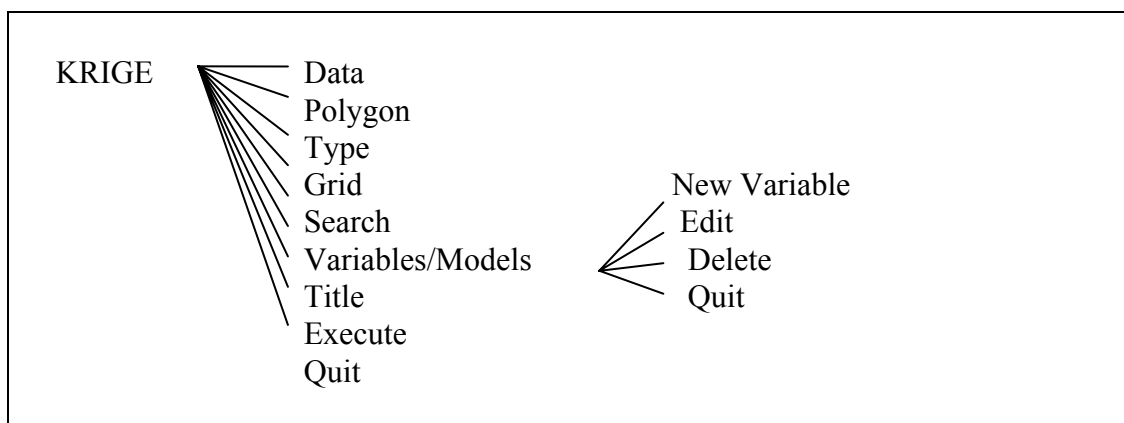


Fig.4.7.Structura subprogramului *Krige*

Sucesiunea opțiunilor care trebuie parcurse pentru interpolarea prin kriging este:

- **Prefix** pentru precizarea identicatorului directorului de lucru în care trebuie să se afle fișierul cu date și în care vor fi salvate rezultatele prelucrării.
- **Read Parameters** este utilizată numai în varianta în care s-a parcurs o sesiune de lucru și parametrii de calcul au fost stocați într-un fișier cu extensia *kpf* (*Kriging Parameters File*).
- **Option/Execute** are subordonate o serie de opțiuni care permit precizarea parametrilor de calcul pentru operațiunea de interpolare prin kriging:

- **Data** pentru precizarea numelui fișierului de date de tip Geo-EAS (cu extensia uzuală ".dat") și a fișierului în care vor fi salvate rezultatele prelucrării (cu extensia uzuală ".grd").

- **Polygon** pentru precizarea numelui fișierului cu conturul poligoanelor în care se face operațiunea de interpolare prin kriging.

Fișierul care conține descrierea poligoanelor are în mod uzual extensia ".pol", este un fișier de tip ASCII și conține (**Fig.4.8**) pe primul rând numărul de poligoane din fișier iar pentru fiecare poligon:

- *codul* pentru precizarea spațiului în care se face interpolarea:
 - 0 – pentru interiorul poligonului;
 - 1 – pentru exteriorul poligonului;
- *numărul de colțuri*;
- *coordonatele fiecărui colț* (în ordinea x, y)

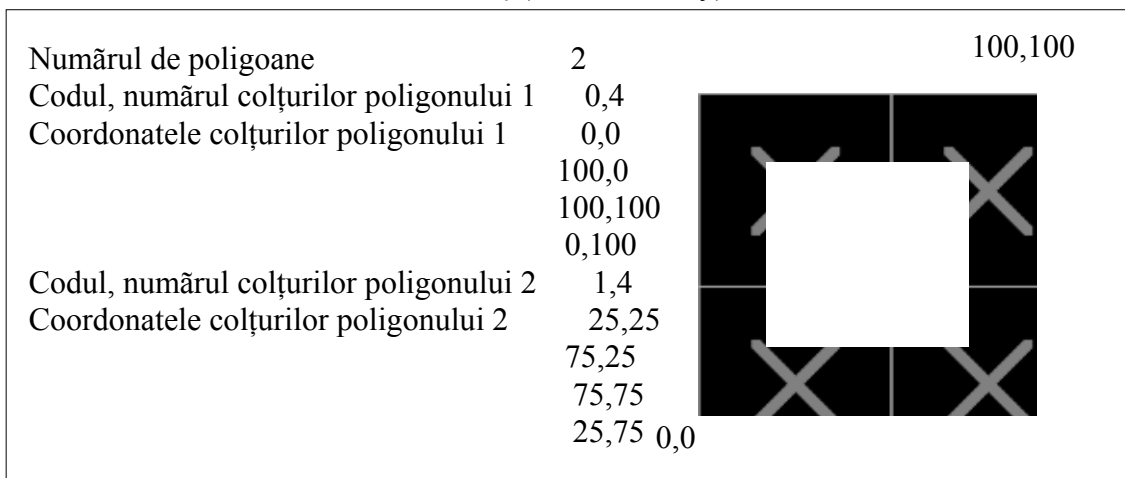


Fig.4.8. Structura fișierului tip "pol" pentru două poligoane pătrate. (interpolarea prin kriging se face numai în zona neagră marcată cu "X"-uri).

- **Type** este opțiunea prin care se alege tipul de kriging utilizat: simplu (*Simple*), ordinar (*Ordinary*), punctual (*Point*) sau zonal cu patru (*Block 2x2*), nouă (*Block 3x3*) și șaisprezece (*Block 4x4*) puncte de discretizare.

- **Grid** este necesar pentru stabilirea densității rețelei de interpolare prin intermediul originii rețelei (*Origin*), echidistanței dintre noduri (*Spacing*) și numărului de noduri (*Number*) pe cele două axe (*X, Y*).

- **Search** delimitează mărimea, forma și structura zonei de calcul pentru fiecare nod din rețea prin: raza mare a zonei de căutare (*Major Radius*), raza minimă de căutare (*Minor Radius*), orientarea axei lungi a elipsei de anizotropie (*Ellipse Angle*), numărul de sectoare ale elipsei (*Sectors*). Pentru creșterea eficienței calculului se poate stabili numărul de sectoare admise fără puncte (*Empty sectors*), numărul maxim de puncte din fiecare sector (*Max in Sector*), precum și cel minim (*Min. to use*).

- **Variables/Models** este pentru precizarea variabilei pentru care se interpolează și a parametrilor modelului de variogramă stabilit cu subprogramul *Vario*.

- **Title** permite scrierea textului care va însoți rezultatele interpolării prin kriging în fișierul care se creează automat de subprogram în faza de execuție a calculelor.

- **Execute** este opțiunea care declanșează calculul de interpolare. Dacă nu au fost introduși toți parametrii sau dacă aceștia nu corespund datelor prelucrate subprogramul oprește calculele. Dacă din alte motive după declanșarea calculelor se dorește întreruperea lor, acest lucru se face prin combinația de taste **Ctrl+Alt+X** după

dezactivarea tastelor **Caps Lock**, **Scroll Lock** și **Num Lock**.

Calculule sunt însoțite de reprezentări grafice și rezultate intermediare ale prelucrării care pot fi activate sau inhibate prin trei taste:

- **Caps Lock** pentru desenarea pe ecran a dispoziției punctelor de observație, poziția punctului de calcul, zona de influență a acestuia și punctele utilizate pentru interpolare;
- **Scroll Lock** pentru afișarea sistemului de kriging;
- **Num Lock** pentru afișarea rezultatelor obținute prin rezolvarea sistemului de kriging.

Continuarea calcululelor după fiecare prezentare se declanșează cu tasta *Q*.

Dacă se dezactivează toate cele trei taste pe ecranul monitorului este prezentată harta rețelei de interpolare, dinamica calcululelor iar la partea inferioară rezultatele pentru fiecare punct al rețelei în ordinea parcurgerii lor.

Rezultatele calcululelor (valoarea estimată și abaterea standard corespunzătoare pentru fiecare punct al rețelei) sunt stocate automat, pe parcursul calcululelor, în fișierul al cărui identificator a fost precizat la opțiunea **Data**.

4.8.SUBPROGRAME UTILITARE Geo-EAS

Subprogramele utilitare ale pachetului Geo-EAS se pot separa în trei categorii:

- a) subprograme utilitare pentru organizarea fișierelor de date: *Dataprep* și *Trans*;
- b) subprograme utilitare pentru reprezentări grafice: *Conrec*, *Scatter*, *Xygraph*, *Postplot*;
- c) subprograme pentru vizualizarea graficelor conținute în fișiere metacode: *View* și *Hpplot*.

4.8.1.DATAPREP

Subprogramul este structurat pe două categorii de operații privitoare la organizarea fișierelor de date. Prima categorie - *DOS Utilities* - permite accesul la opțiuni tip DOS iar cealaltă - *File Operations* - include utilități pentru crearea sau modificarea fișierelor de date de tip Geo-EAS. Operațiile cu fișiere de date se fac cu ajutorul unui fișier temporar numit *Scratch File*.

Opțiunile din nivelul **Dos Utilities** organizate pe un ecran standard (**Fig.4.9**) sunt:

- **Directory** pentru afișarea conținutului unui director al cărui nume trebuie precizat;
- **Print** pentru scrierea la imprimantă a conținutului fișierului specificat;
- **List** pentru listarea pe monitor a conținutului fișierului specificat;
- **Copy** pentru copierea fișierelor;
- **Rename** pentru redenumirea fișierelor;

- **Delete** pentru ștergerea fișierelor;
- **DOS Command** pentru ieșirea temporară din program și plasarea în "prompter DOS". Revenirea în program se face cu comanda EXIT.

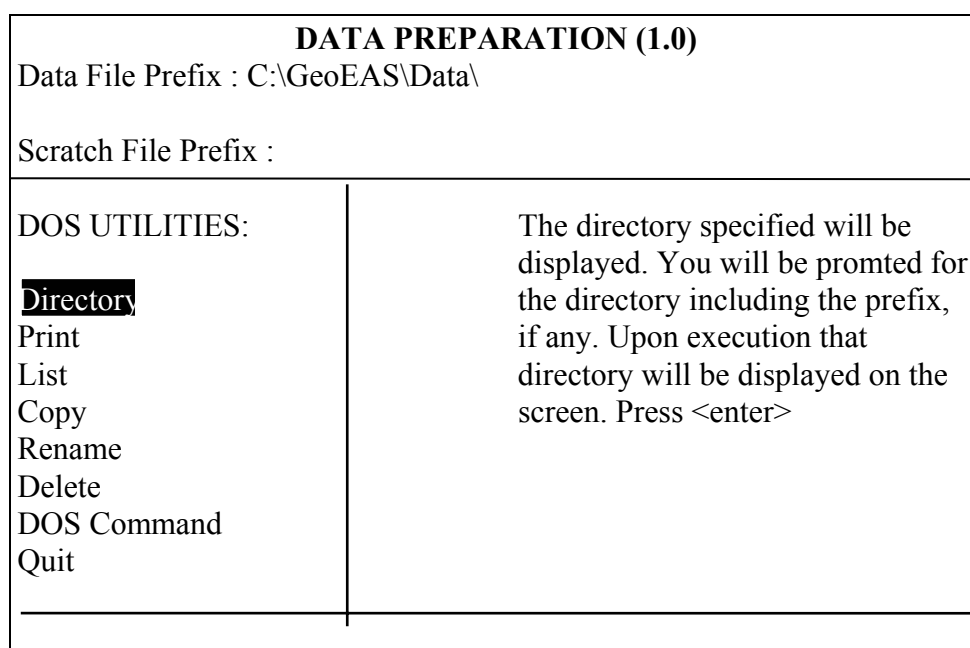


Fig. 4.9. Ecranul programului Dataprep DOS Utilities

Opțiunile din nivelul **File Operations** organizate într-un ecran similar cu cel din **Fig.4.9** realizează modificări în fișierele de date prin:

- **Append** pentru crearea unui fișier nou din conținuturile a două fișiere de tip Geo-EAS. Când cele două fișiere (Fișierul 1 și Fișierul 2; **Fig.4.10**) care se combină conțin variabile diferite (v1, v2), în fișierul rezultat (Fișierul 1+2) spațiile fără valori sunt ocupate cu codul numeric pentru valori absente (1E+11).

Fișierul 1		Fișierul 2		Fișierul 1+2	
3		3		4	
x	m	x	m	x	m
y	m	y	m	y	m
v1	ppm	v2	ppm	v1	ppm
20, 30, 300		20, 30, 280		v2	ppm
45, 60, 250		45, 60, 115		20, 30, 300, 280	
		30, 85, 80		45, 60, 250, 115	
				30, 85, 1E+11, 80	

Fig.4.10. Modalitatea de concatenare a două fișiere de tip Geo-EAS cu Append

- **Column Extract** pentru extragerea valorilor de pe o coloană în funcție de condiții numerice de tipul:
 - *mai mic decât* (.LT.; engl. "less than"), *mai mic sau egal cu* (.LE.; engl. "less or equal to");
 - *mai mare decât* (.GT.; engl. "greater than"), *mai mare sau egal cu* (.GE.; engl. "greater than or equal to");
 - *egal cu* (.EQ.; engl. "equal to"), *diferit de* (.NE.; engl. "not equal to").
- **Row Extract** pentru extragerea valorilor de pe un rând în funcție de anumite

condiții numerice (similare cu cele de la *Column Extract*);

- **Compress** realizează compactarea fișierelor de tip Geo-EAS prin ștergerea rândurilor care se repetă;
- **ID Variable** pentru crearea unei variabile de tip număr curent care se adaugă celor existente în fișierul sursă;
- **Merge** pentru concatenarea a două fișiere care au aceleași puncte de observație și variabile diferite;
- **Report** pentru prezentarea datelor dintr-un fișier de tip Geo-EAS sub forma unui raport tip DATABASE;
- **Sort** pentru ordonarea valorilor într-un fișier după una din variabilele selectate.

4.8.2. TRANS

Subprogramul *Trans* este utilizat pentru modificarea unui fișier de date sau crearea unuia nou. Structura subprogramului are opțiunile organizate pe șase niveluri:

- **Prefix** pentru precizarea identificatorului directorului de lucru;
- **Data** pentru precizarea identificatorului fișierului cu date;
- **Create** pentru crearea unei variabile noi (*New Variable*) sau modificarea uneia vechi (*Old Variable*);

- **New Variable** pentru crearea unei noi variabile prin operații de două tipuri:

- **Unary Operation** care operează asupra valorilor unei singure variabile cu diferite tipuri de prelucrări selectate cu "space" dintr-un repertoriu limitat:

- + adunare;
- - scădere ;
- * înmulțire;
- / împărțire;
- ** ridicarea la putere;
- *sqrt* radical din variabila aleasă;
- *log* logaritm zecimal din variabila aleasă;
- *ln* logaritm natural din variabila aleasă;
- *truncate* reținerea părții întregi din variabila aleasă;
- *exp* = $(e)^x$, unde $e = 2.71828$ și x este variabila aleasă;
- *rank* - determină rangul valorilor variabilei selectate și creează o nouă variabilă cu aceste valori.

- **Constant** permite introducerea constantelor numerice.
- **Variable** permite selectarea variabilei cu care se operează.
- **Execute** realizează prelucrarea proiectată.
- **Binary Operation** operează similar cu **Unary Operation** prelucrând simultan două variabile.

- **Indicator Transform** operează asupra valorilor unei variabile prin intermediul unei valori de prag (*threshold value*) introdusă prin intermediul opțiunii **Cutoff**. Rezultatul este un șir format din 0 și 1: 1 (unu) pentru valori ale variabilei mai mari ca valoarea de prag și 0

(zero) pentru valorile mai mici ca valoarea de prag.

- **V_Delete** permite ștergerea unei variabile din fișier, variabila fiind selectată cu "space".
- **Save** permite salvarea rezultatelor transformărilor realizate într-un fișier al cărui identificator trebuie precizat.
- **Quit** este opțiunea de ieșire din subprogram.

4.8.3. SCATTER

Subprogramul *Scatter* realizează reprezentarea grafică a două variabile într-o diagramă rectangulară. Opțional subprogramul calculează parametrii corelației liniare între cele două variabile:

- *ordonata la origine* (eng.: intercept);
- *panta* (eng.: slope);
- *coeficientul Pearson* (eng.: Corell.coef.)

Structura subprogramului este organizată pe un singur nivel cu opțiunile:

- **Prefix** pentru introducerea identificatorului directorului de lucru;
- **Data** pentru precizarea identificatorului fișierului de date;
- **Variable** pentru selectarea variabilelor;
- **Option** pentru selectarea calculului parametrilor corelației liniare;
- **Execute** pentru reprezentarea diagramei rectangulare;
- **Quit** pentru ieșirea din subprogram.

Subprogramul *Scatter* este utilizat în faza inițială a prelucrărilor geostatistice pentru construirea hărților cu poziția punctelor de probare sau pentru identificarea corelațiilor de tip linear între variabilele studiate.

4.8.4. XYGRAPH

Subprogramul *Xygraph* realizează reprezentarea punctuală a maximum șase variabile într-o diagramă rectangulară utilizând o structurare pe trei niveluri a opțiunilor:

- **Prefix** pentru introducerea identificatorului directorului de lucru;
- **Read Parameters** pentru citirea parametrilor de lucru stocați într-o sesiune anterioară.
- **Option/Execute** pentru trecerea la nivelul al doilea de lucru în care se introduc parametrii de lucru utilizând opțiunile:
 - **Data** pentru introducerea identificatorului fișierului cu date;
 - **Variables** pentru selectarea variabilelor care vor fi reprezentate;
 - **Symbol/Line** pentru alegerea simbolurilor și tipurilor de linii utilizate în reprezentarea grafică;
 - **Regression** pentru selectarea calculului parametrilor corelației liniare;
 - **Legend** pentru definirea legendei;
 - **Graph Option** pentru definitivarea graficului utilizând:
 - **Axis Parameters** pentru aspectul axelor graficului;

- **Tick Parameters** pentru scara grafică a axelor;
- **Titles/Labels** pentru textul graficului și al etichetelor punctelor;
- **Reset** pentru revenirea la parametrii implicați ai graficului.
- **Execute** pentru realizarea graficului cu parametrii stabiliți;
- **Save Parameters** pentru stocarea parametrilor de lucru.

4.8.5. POSTPLOT

Subprogramul Postplot realizează hărți punctuale pentru date stocate în fișiere de tip Geo-EAS. Structura ierarhică a subprogramului este similară cu cea a subprogramului *Xygraph* iar singurele opțiuni particulare sunt la nivelul *Options* pentru:

- plasarea valorii lângă poziția punctului de probare (*Include Values*);
- scara de reprezentare a simbolurilor (*Size, Scale Factor*);
- stabilirea numărului de caracteristici zecimale ale valorilor (*Decimal*);
- stabilirea tipului de simboluri (*Include Symbole*).

4.8.6. CONREC

Subprogramul *Conrec* realizează hărți cu izolinii pe baza rețelelor de interpolare create prin kriging și stocate în fișiere de tip Geo-EAS. Rețelele create prin kriging sunt formate din valori distribuite într-o rețea rectangulară egal distanțate pe direcția OX și OY.

Structura este organizată pe trei niveluri și are opțiuni identice cu *Xygraph* până la nivelul opțiunilor pentru trasarea izoliniilor. Acestea sunt plasate pe nivelul al treilea la *Contour Options* :

- **New Levels** permite alegerea unui set de izolinii diferențiate printr-o echidistanță aleasă.
- **Edit Levels** permite editarea unui set de izolinii pentru orice valoare cuprinsă între valoarea minimă și cea maximă a variabilei studiate.
- **Labeling** permite plasarea valorilor pe izolinii.
- **Dash Pattern** permite alegerea aspectului grafic al izoliniilor.
- **Annotation** permite plasarea textelor pe axe.
- **Spline** permite alegerea gradului de "netezire" a izoliniilor.

Execute realizează prelucrările pe baza parametrilor introduși și a unui fișier intermediar de tip *metacode.met*.

Save Parameters permite stocarea parametrilor de construire a hărții cu izolinii într-un fișier de tip **.cpf (Contour Parameters File)*.

4.8.7. VIEW

Subprogramul *View* permite afișarea pe ecranul monitorului a graficelor stocate în fișiere de tip metacode. Opțiunile sunt organizate într-o structură cu un singur nivel și sunt comune tuturor subprogramelor:

- **Prefix** pentru precizarea identificatorului directorului de lucru;
- **File** pentru precizarea identificatorului fișierului de tip metacode care conține graficul a cărui vizualizare se realizează;
- **Scale** pentru stabilirea raportului scârilor de pe cele două axe;
- **Execute** pentru realizarea reprezentării;
- **Quit** pentru părăsirea programului.

4.8.8.HP PLOT

Subprogramul *Hpplot* transformă instrucțiunile de plotare independente de perifericul grafic într-un fișier de comenzi HPGL. Acest fișier poate fi transmis unui ploter HP prin rularea fișierului de comenzi HPSETUP.BAT și a comenzii PRINT din prompter DOS. Structura subprogramului este similară celei a subprogramului *View*.

COMENTARIU

*Nu uitați ca **programelele sunt instrumentele** care permit aplicarea unei metodologii de prelucrare a informațiilor.*

Ca să nu discreditați programul Geo-EAS prin utilizarea lui incorectă trebuie:

- *să știți foarte clar ce vreți să obțineți prin prelucrarea datelor;*
- *să alegeți metodologia de prelucrare în funcție de calitatea și cantitatea datelor disponibile;*
- *să cunoașteți în detaliu fiecare etapă a metodologiei de prelucrare a datelor;*
- *să nu lăsați programul să lucreze cu parametrii implicați ci să-i stabiliți pe baza analizei datelor prelucrate*

*Nu încercați să “**ghiciți**” metodologia de prelucrare a datelor prin utilizarea programului înaintea înțelegerii acesteia!*

Dacă rezultatul estimării unei structuri spațiale nu este credibil nu acuzați programul Geo-EAS! El execută ce i se comandă. Aveți grijă ce-i cereți să facă!