

1.1 Modele probabiliste utilizate în proiectarea rețelelor de monitorizare

Mărimea fundamentală implicată în modelele probabiliste de proiectare a rețelelor de monitorizare este **entropia** a cărei minimizare crește gradul de cunoaștere al proceselor studiate.

Entropia reprezintă reducerea incertitudinii unui eveniment care se produce cu probabilitatea p :

$$H = \log(p) \quad (3.1)$$

Entropia este **aditivă** când este aplicată intersecției a două evenimente independente:

$$H = p \cdot (-\log p) + (1-p) [-\log(1-p)] \quad (3.2)$$

Cele două evenimente pot fi două distribuții punctuale cu probabilitățile p și $(1-p)$.

Această definiție poate fi extinsă la distribuția a mai multor puncte (ex.: punctele de probare ale rețelei de probare/monitoring).

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).

Pentru un **mediu continuu** nu este posibil de găsit o echivalență riguroasă. 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 analogă sub forma:

$$H(V) = E[-\log f(V)] \quad (3.3)$$

Analogia cu mediul discret nu este satisfăcătoare deoarece nu este invariantă la transformările realizate asupra variabilei V , f (funcția densității de probabilitate) fiind exprimată în unități de probabilitate raportate la unități de V .

Pentru uniformizare dimensională este propusă (Jaynes, 1968) forma:

$$H(V) = E \left[-\frac{\log f(V)}{m(V)} \right] \quad (3.4)$$

în care $m(V)$ este o măsură a "ignorantei 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.

1.1.1 Entropia seriilor de timp

Analiza **seriilor de timp**, din punct de vedere al posibilităților de prognoză poate fi abordată pe baza diverselor metode probabiliste.

În cazul evoluției complexe a parametrilor stabilității versanților identificarea unor factori determinanți și a unor relații cauzale este dificilă, cel mai adaptabil model probabilist fiind cel al **lanțurilor Markov**, model în care starea sistemului la un moment t se acceptă a fi determinată numai de starea sistemului la momentul $(t-1)$.

Metoda implică definirea unor **variabile aleatoare dinamice** care să descrie cu acuratețe evoluția în timp a parametrului studiat.

Aceste variabile aleatoare dinamice asociate cu fiecare stație de monitorizare și fiecare parametru analizat sunt de tipul:

$$X: \begin{pmatrix} t_1 & t_2 & t_3 & \dots & t_i & \dots & t_{n-1} & t_n \\ p_1 & p_2 & p_3 & \dots & p_i & \dots & p_{n-1} & p_n \end{pmatrix} \quad (3.5)$$

in care

t_1, t_2, \dots, t_n - stări în evoluția parametrului studiat, înregistrate în momente succesive din evoluția parametrului;

p_1, p_2, \dots, p_n - probabilitățile stărilor înregistrate în evoluția parametrului.

Pentru analiza evoluției în timp a parametrilor, fiecărui punct/stație de monitorizare ii sunt asociate atâtea variabile aleatoare câți parametrii sunt determinați (ex.: precipitații, nivel hidrostatic, umiditate, panta terenului etc.).

Stările caracteristice alese pentru evoluția parametrilor studiați pot fi:

- a-creșterea valorii parametrului;
- b-mentinerea constantă a valorii parametrului;
- c-descreșterea valorii parametrului.

Acest tip de variabile aleatoare dinamice permit descrierea evoluției valorii parametrului doar prin intermediul corelației stării la momentul t_i cu starea la momentul t_{i-1} tipică unui **lanț Markov**.

Descrierea completă a evoluției unui parametru se realizează pe baza matricii de tranziție de tipul:

$$M: \begin{pmatrix} P_{11} & P_{12} & P_{1i} & P_{1n} \\ P_{21} & P_{22} & P_{2i} & P_{2n} \\ \dots & \dots & \dots & \dots \\ P_{n1} & P_{n2} & P_{ni} & P_{nn} \end{pmatrix} \quad (3.6)$$

in care

P_{ij} - probabilitatea de tranziție din starea "i" în starea "j".

Complexitatea evoluției sistemului poate fi detaliată prin multiplicarea stărilor analizate (cuantificându-se și mărimea variației într-un număr de clase adecvat precizie de măsurare) evaluarea acestora realizându-se prin intermediul entropiei calculate cu relația:

$$H_i = -\sum_1^n P_i \times \log_2(P_i); [bit] \quad (3.7)$$

in care

P_i - probabilitatea de atingere a starii i ;

H_i - entropia procesului în care fiecare stare este considerată independentă de celelalte.

Entropia pentru fiecare parametru și punct de monitorizare este calculată în acord cu mărimea erorilor de măsură. Valoarea maximă a entropiei independente (H_i) este obținută din relația:

$$H_{i_{max}} = -\log_2 s, [bit] \quad (3.8)$$

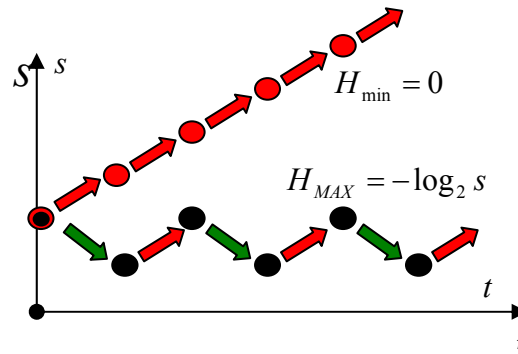


Fig.3.1. Entropiile extreme ale unei serii de timp.

și corespunde situației în care toate cele s stări ale sistemului sunt egal posibile (**Fig.1**). Aceasta este starea de maximă incertitudine privind evoluția și prognoza pe termen lung a sistemului. Pentru un sistem cu trei stări (a, b, c) entropia maximă este :

$$H_{MAX} = \log_2(3) = 1,58 \quad (3.9)$$

Este evident că stările sistemului nu sunt independente motiv pentru care pe baza modelului markovian, utilizând matricea de tranziție a procesului se calculează entropia de corelație (de dependență; H_d) cu relația:

$$H_d = -\sum_1^n P_i \sum_1^n P_{ij} \log_2 P_{ij} \quad (3.10)$$

Entropia de dependenta este cu atât mai mică cu cât corelația între stări este mai strânsă.

Pe baza valorilor entropiilor H_d și H_i , este posibilă zonarea incertitudinii prognozei pentru evoluția parametrului studiat și ca rezultat al acestuia se separa zonele :

- cu evoluție previzibilă;
- cu evoluție imprevizibilă (situație care implică o creștere a numărului de valori măsurate, prin reducerea intervalului de timp dintre măsuratori.

Optimizarea regimului de observații în fiecare stație de monitorizare pentru fiecare proprietate studiată se bazează pe minimizarea entropiei de dependență (H_d).

1.1.2 Entropia rețelei de monitorizare

Abordarea studiului entropiei unei rețele de monitorizare pornște 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 măsuratori candidează pentru funcția de **stații permanente** în rețeaua de monitorizare. Densitatea stațiilor de măsurare este limitată din considerente practice (costul execuției stației de monitorizare și al măsurătorilor executate). Minimizarea entropiei rețelei de monitorizare vizează stabilirea **numarului minim de stații permanente** din rețea.

Fie X vectorul aleator ale cărui coordonate reprezintă valorile câmpului spațial în diferitele puncte ale rețelei de monitorizare ordonate după un criteriu rezonabil.

Stațiile permanente ale rețelei de monitoring vor fi selectate din aceste puncte disponibile iar problema proiectantului este identificarea acelor care trebuie măsurate.

Pentru această operațiune partiția vectorului X poate fi reprezentată sub forma:

$$X = (U, G)$$

în care

U reprezintă valorile variabilei în punctele nemonitorizate (stații nepermanente);

G - reprezintă valorile în puncte monitorizate.

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. Presupunind că $m(G)$ este specificat atunci:

$$m(U|G) = \frac{m(U, G)}{m(G)} \quad (3.11)$$

Pentru un anumit proces stocastic, incertitudinea rețelei de monitoring se poate exprima prin entropia variabilei X :

$$H(X) = H(U|G) \quad (3.12)$$

care se poate descompune în :

$$H(U, G) = H(U|G) + H(G) \quad (3.13)$$

în care

$$H(U|G) = E \left[- \frac{\log f(U|G)}{m(U|G)} \right] \quad (3.14)$$

Pentru a interpreta această decompoziție, presupunem ca valorile G au fost obținute și în aceste condiții incertitudinea asupra valorilor U este :

$$H(U) = E \left[- \frac{\log f(U|G)}{m(U|G)} \Big| G \right] \quad (3.15)$$

Deoarece G nu este cert, incertitudinea relativă la U este primul termen în decompoziția entropiei $H(X) = H(U, G)$:

$$H(U|G) = E \left\{ E \left[- \frac{\log f(U|G)}{m(U|G)} \Big| G \right] \right\} \quad (3.16)$$

deci

$$H(X) = E \left\{ E \left[- \frac{\log f(U|G)}{m(U|G)} \Big| G \right] \right\} + E \left[- \frac{\log f(G)}{m(G)} \right] \quad (3.17)$$

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 și astfel să se maximizeze beneficiul măsurătorilor executate.

Este ușor de arătat că:

$$H(U|G) = E \left[- \frac{\log(U|G)}{m(U|G)} \right] < E \left[- \frac{\log f(U)}{m(U)} \right] = H(U) \quad (3.18)$$

În aceste condiții obținând G prin măsuratori niciodată nu va avea loc o creștere a incertitudinii asupra partitiei U . Presupunând $m(U|G) = m(U)$ rezultă:

$$H(U|G) = H(U) - I(U; G) \quad (3.19)$$

în care informația mutuală în U și G este strict pozitivă, simetrică și definită de relația

$$I(U; G) = E \left\{ \log \left[\frac{f(U, G)}{f(U) \cdot f(G)} \right] \right\} \quad (3.20)$$

Obiectivul minimizării entropiei se realizează în contextul instalării unei noi rețele sau al extinderii uneia existente.

În ambele cazuri efectul unui număr restrins de puncte măsurate suplimentar repartizate pe întreaga suprafață monitorizată are un efect 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 :

$$H(U) - H(U|G) = I(U; G) = I \quad (3.21)$$

Posibilitatea includerii în U a valorilor cu înaltă incertitudine 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 reduce stațiile de observație cu cea mai mică pierdere de informație posibilă.

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. Se poate încorpora în mod direct această incertitudine a modelului în incertitudinea asupra distribuției procesului analizat sub forma:

$$H(U, G, q) = H(U, G|q) + H(q) \quad (3.22)$$

$H(q)$ - incertitudinea asupra parametrilor modelului stocastic utilizat;

$H(U, G, q)$ - incertitudinea totală a variabilei X .

În studiul rețelelor 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 acestora.

Entropiile instrumentale și alte forme de erori de măsurare trebuie și ele luate în considerare în proiectarea rețelelor 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:

$$H(U, G, q) - H(U, G, q|D) = H(G) - H(G|D) \quad (3.23)$$

cu măsura "ingnoranței complete":

$$m(U, G, q|D) = m(U, q|G) \cdot m(G|D) \quad (3.24)$$

Atâta timp cât $H(U, G, q)$ este constantă și independentă de alegerea procesului de măsurare a variabilei G și de modul de partiționare (U, G) se deduce că $H(U, G, q|D)$ poate fi minimizată prin combinația modului de alegere a procesului de măsurare G și de partiționare al variabilei X .

În general $f(U, q|G, D) = f(U, q|G)$ și dacă $m(U, G, q|D) = m(U, q|G) \cdot m(G|D)$ rezultă că:

$$H(U, G, q|D) = H(U, q|G) + H(G|D) \quad (3.25)$$

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 analiza. 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 $m(D|G) = \frac{m(G|D) \cdot m(D)}{m(G)}$ atunci

$$H(G|D) = H(G) - [H(D) - H(D|G)] \quad (3.26)$$

relație care exprimă faptul că o bună cunoaștere a variabilei G reduce substanțial incertitudinea asupra distribuției variabilei X .

Transformate în mod adecvat, multe serii de valori ale parametrilor monitorizați au distribuții care pot fi aproximare printr-o distribuție gaussiană.

În contextul noțiunilor introduse, variabila X este vectorul valorilor parametrilor măsurați în stațiile unei rețele de monitoring la un moment dat. Coloanele matricii de date $D(X_1, X_2, \dots, X_n)$ sunt presupuse apriori a fi independente. O altă ipoteză plauzibilă este aceea ca $\{X_i\}$ are o distribuție normală multivariată cu un vector mediu m și matricea de covarianță Σ , distribuție care poate fi exprimată simbolic:

$$X_i|m, \Sigma \approx indN_p(m, \Sigma) \quad (3.27)$$

p - numărul de puncte de observație din rețea;

n - numărul de momente în care s-au executat măsurători ($= 1, \dots, n$)

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 $(\Sigma_{GG}, \Sigma_{U|G}, t)$ conform relațiilor:

$$\Sigma_{GG} = \text{cov}(G, G)^Y = E[G - E(G)][G - E(G)]^T \quad (3.28)$$

$$\Sigma_{U|G} = \Sigma_{UU} - \Sigma_{UG} \cdot \Sigma_{GG}^{-1} \cdot \Sigma_{GU} \quad (3.29)$$

Matricea $t = \Sigma_{U|G} \cdot \Sigma_{GG}^{-1}$ este panta predictorului linear optimal al partiției U prin G (adică $E(U) + t[G - E(G)]$).

Această transformare este realizată prin decompoziția Bartlett $\Sigma = T\Delta T^T$ în care

$$\Delta = \begin{bmatrix} \Sigma_{U|G} & 0 \\ 0 & \Sigma_{GG} \end{bmatrix} \quad T = \begin{bmatrix} l & \tau \\ 0 & l \end{bmatrix} \quad (3.30)$$

$$\Sigma^{-1} = (T^T)^{-1} \Delta^{-1} T^{-1} \quad (3.32)$$

$$\Delta^{-1} = \begin{bmatrix} \Sigma_{U|G}^{-1} & 0 \\ 0 & \Sigma_{GG}^{-1} \end{bmatrix} \quad T^{-1} = \begin{bmatrix} l & -\tau \\ 0 & l \end{bmatrix} \quad (3.33)$$

Pentru un set de date $d(x_1, x_2, \dots, x_n)$ descompunerea entropiei se face în trei componente:

$$E \left[-\frac{\log f(X, \mu, \Sigma|d)}{m(X, \mu, \Sigma|d)} \middle| d \right] = HU + HMODEL + HG \quad (3.34)$$

HU reprezintă incertitudinea reziduală apriori a partiției U după utilizarea partiției G pentru evaluarea valorilor partiției U pe baza modelului linear:

$$HU = E[-\log f(U|G, \mu, \Sigma)|d] \quad (3.35)$$

HG reprezintă incertitudinea asupra partiției G eliminată prin monitorizare și necesar de maximizat printr-o partiție optimă a variabilei X :

$$HG = E[-\log f(G|d)|d] \quad (3.36)$$

HMODEL reprezintă incertitudinea reziduală asupra partiției U , mediei și matricii de covarianță corespunzătoare unei anumite partiții G măsurate:

$$HMODEL = E\left[-\frac{\log f(\mu, \Sigma|G.d)}{m(\mu, \Sigma)}|d\right] \quad (3.37)$$

in care Σ și m sunt parametrii ai distribuției Wishart (Anderson, 1984 p268 pentru semnificații și definiții).

Optimizarea rețelei se fundamentează pe minimizarea componentelor HU și $HMODEL$.

1.1.3 Proiectarea rețelelor de monitoring

Adoptarea drept criteriu de proiectare a rețelelor de monitoring a reducerii entropiei totale are implicații în politica proiectării și în implementare.

Criteriul reducerii entropiei este bazat pe maximizarea reducerii incertitudinii asupra partiției U calculată prin măsurătorile realizate în stațiile partiției G . Acest criteriu are ca singur obiectiv predicția partiției U din G pe baza datelor obținute prin monitorizarea stațiilor partiției G .

Sub aspect practic el permite predicția valorilor în stațiile nemonitorizate și estimarea parametrilor distribuției multivariate a unui vector aleator format din măsurătorile executate în toate locațiile. Ambele obiective se realizează prin minimizarea sumei incertitudinilor măsurate prin entropia lor.

Minimizarea sumei acestor două incertitudini poate fi echivalată într-o primă variantă (Caselton și Husain, 1980) cu maximizarea incertitudinii transferate partiției G , eliminată prin monitorizare stațiilor din G .

Este important de sesizat că numai maximizarea componentei HG este un obiectiv prea îngust și că este necesar să fie completat cu considerații suplimentare asupra partiției U . Această observație a condus la ideea maximizării informației mutuale ($I(U, G)$; Caselton și Zidek, 1984) și echivalării maximizării componentei HG cu minimizarea sumei componentelor $HU + HMODEL$.

În practica actuală, cele două criterii operaționale utilizate sunt:

- maximizarea informației mutuale: $I(U, G)$;
- minimizarea sumei incertitudinilor: $HU + HMODEL$.

Trei situații principale apar în proiectarea rețelelor de monitoring:

- initializarea unei rețele de monitoring;
- extinderea unei rețele existente;
- eliminarea unor stații dintr-o rețea existentă.

1.1.3.1 Initializarea unei rețele de monitoring

Inițializarea unei rețele 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**. În practică realizarea unei astfel de rețele este împiedicată de:

- gradul de accesibilitate la punctele proiectate în teren;
- asigurarea securității stațiilor de monitoring;
- dorința de a minimiza costul execuției;
- dorința de a realiza un număr maxim de stații dintr-un buget fix.

Când cunoașterea distribuției spațiale este vagă, o bază obiectivă și rațională pentru proiectarea unei rețele de monitoring necesită anumite date colectate înaintea unei **rețele permanente**. Numai dacă nu se pot obține prin teledetecție 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.

Coroborarea 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 posibilă de funcționare a rețelelor temporare.

Criteriul entropiei se aplică în două alternative: cel al maximizării informației mutuale și cel al minimizării incertitudinii $HU + HMODEL$.

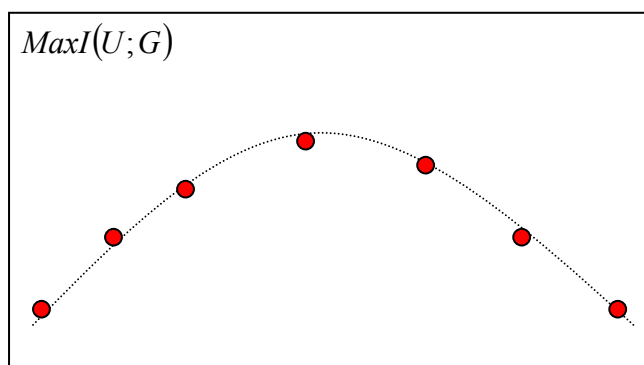
1.1.3.1.1 Maximizarea informației mutuale $I(U, G)$

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 (**Fig.3.2**).

Acest criteriu este eficient pentru rețele cu un număr de stații (g) mult mai



1 Număr de stații g în rețea

Fig.3.2. Corelația între valoarea maximă a informației mutuale și numărul de stații din rețeaua de monitorizare

mic decât cel al **stațiilor temporare** (p). Dacă nu sunt impuse condiții în alegerea locațiilor atunci curba $MaxI(U;G) = f(p)$ este simetrică. Nici o indicație asupra numărului optim de stații g , nu se poate extrage din acesta curbă.

1.1.3.1.2 Minimizarea incertitudinii $HU + HMODEL$

Acest criteriu plasează pe ultimul loc rolul informațiilor obținute din rețeaua temporară utilizată pentru proiectarea rețelei temporare deoarece echivalentul lui este maximizarea parametrului $|\Psi_{GG}|$.

El trebuie evaluat numai pentru un set de locații selectat 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 (**Fig. 3.3**).

Posibilitatea alegerii numărului optim de stații în rețeaua de monitoring este dată de reducerea la minimum a incertitudinii $HU + HMODEL$ la adăugarea ultimei locații.

Minimum reducerii este greu de absolutizat. Mai riguros această reducere poate fi comparată pentru introducerea diferitelor stații în rețeaua de monitoring.

1.1.3.2 Extinderea unei rețele existente

Situația este similară cu inițializarea unei rețele diferențiale 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 $I(U, G)$.

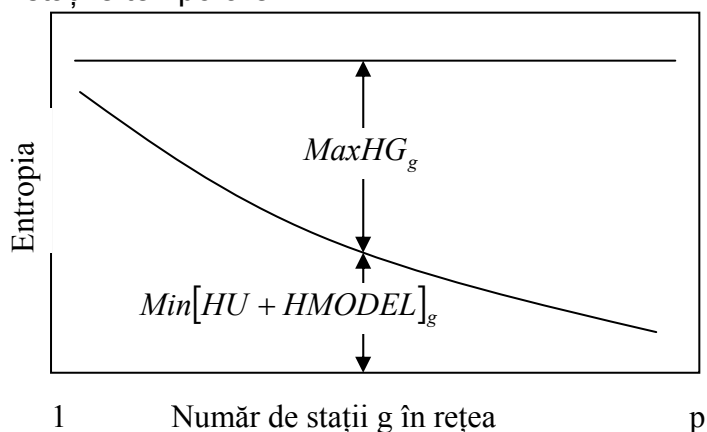


Fig.3.3. Corelația între HG , $HU + HMODEL$ și numărul de stații din rețeaua de monitorizare

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.

1.1.3.3 Eliminarea unor stații din rețeaua existentă

Această problemă poate fi abordată în același mod ca initializarea rețelei, excepție făcând faptul că stațiile existente permit obținerea evoluției parametrilor pe o lungă perioadă în toate stațiile. Este eliminat în acest caz dezavantajul unei informații subiective. Singura situație care poate deranja este aceea în care cu toate că dispunem de un număr suficient de stații, istoricul măsurătorilor să nu fie adecvat evaluării corecte a distribuției parametrilor.

Eficiența metodei evaluării rețelelor de monitoring pe baza entropiei este determinată în principal de calitatea datelor (distribuție spațială și istoricul măsurătorilor). Ca o generalizare, metoda minimizării entropiei poate fi extinsă și în studiul impactului structurii temporare a seriilor de măsurători asupra clasificării stațiilor rețelelor de monitoring.

1.2 Estimarea topo-probabilistă a parametrilor

Estimarea topo-probabilistă a parametrilor are ca obiective:

- evitarea supraestimării sau subestimării valorilor caracteristice ale parametrilor geotehnici;
- reducerea la minimum a erorilor de estimare a valorilor medii și a distribuției spațiale a acestora.

Realizarea acestor două obiective se bazează pe trei etape de prelucrare distincte:

- analiza variabilității globale a parametrilor geotehnici;
- analiza variabilității spațiale a caracteristicilor litologice și parametrice;
- estimarea punctuală

1.2.1 Analiza variabilității globale

Analiza variabilității *globale* a caracteristicilor geotehnice 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 acestora* 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*, 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 definatorii 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.Bomboc, 1979).

1.2.1.1 Normalizarea repartiției selecțiilor de date

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):

$$t_i = T(v_i)$$

(3.38)

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 normate* și este cunoscută sub numele de *normalizare redusă*.

Normalizarea redusă a distribuțiilor în *câmpul valorilor normate* 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 :

$$u_i = \frac{v_i - m}{s} \quad (3.39)$$

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:

$$p_1 = p_2 = \dots = p_n = \frac{1}{n} \quad (3.40)$$

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 :

$$\sum_{i=1}^n p_i = 1 \quad (3.41)$$

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.

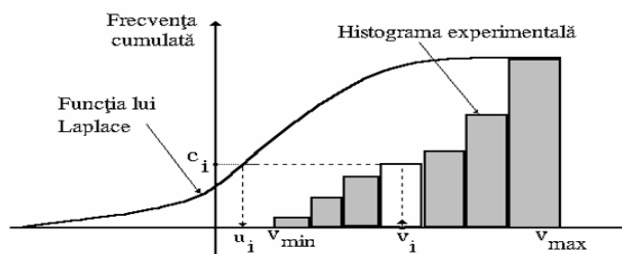


Fig.3.4.Principiul normalizării valorilor.

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.3.4)*.

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:

$$u_i = G^{-1}(c_i) \quad (3.42)$$

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 normale 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.

1.2.2 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 intervareograma) 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 celei mai probabile legi de variație spațială a caracteristicii studiate.

1.2.2.1 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 ($|\vec{h}_{ij}|$) pentru punctele P_i și P_j din **Fig.3.5.**;
- orientarea acestui segment (θ din Fig. 1.17), măsurată prin unghiul dintre direcția axei abscisei și a segmentului care unește cele două puncte.

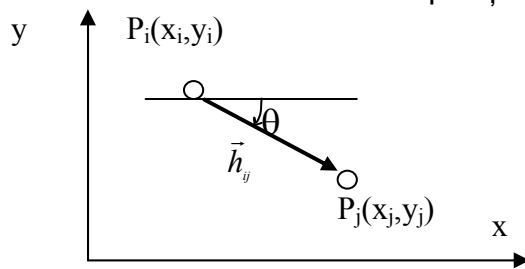


Fig.3.5. Vectorul de poziție \vec{h}_{ij}

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 ($v_i(P), v_j(P + \vec{h}_{ij})$) se reprezintă prin puncte într-un sistem de referință rectangular (**Fig.3.6**).

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

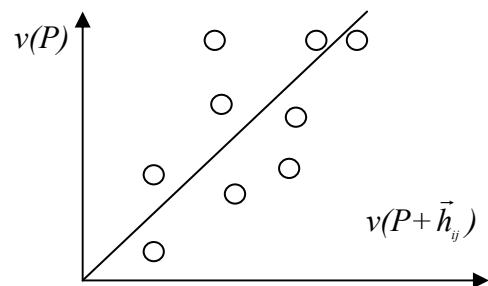


Fig.3.6. Diagrama de continuitate

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 (\vec{h}_y).

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 abatereii standard (s) corespunzătoare. Existența corelației între valoarea medie și abaterea standard este cunoscută sub denumirea de efect de proporționalitate. Efectul de proporționalitate directă (Fig.2.19a) 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;
- 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.2.19b), 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.

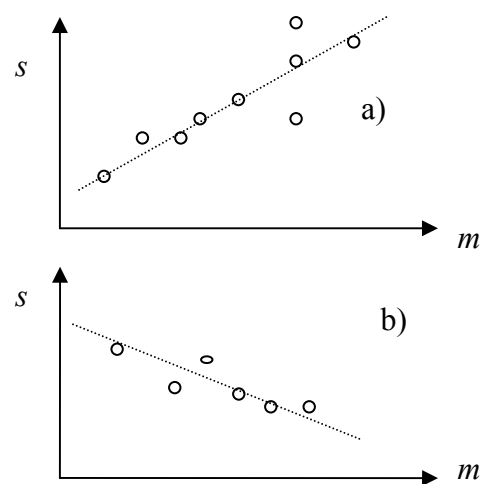


Fig.3.7. Diagrame de variabilitate cu efect de proporționalitate directă (a) și inversă (b).

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.

1.2.2.2 Analiza parametrică a datelor

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 ($\square(h)$) și funcția de variogramă ($\square(h)$).

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 acestora 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:

$$c(h) = \frac{1}{N(h)} \sum_{(i,j)_{|h_{ij}=h}}^{N(h)} v_i v_j - m_{-h} m_{+h} \quad (3.43)$$

în care:

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

$(i,j)_{|h_{ij}=h}$ - perechea de puncte (p_i, 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 $|\vec{h}|$.

$$m_{-h} = \frac{1}{N(h)} \sum_{(i,j)_{|h_{ij}=h}}^{N(h)} v_i \quad (3.44)$$

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

$$m_{+h} = \frac{1}{N(h)} \sum_{(i,j)_{|h_{ij}=h}}^{N(h)} v_j \quad (3.45)$$

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

$$\rho(h) = \frac{c(h)}{\sigma_{-h} \sigma_{+h}} \quad (3.46)$$

în care:

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

$$\sigma_{-h}^2 = \frac{1}{N(h)} \sum_{i,j \in h} v_i^2 - m_{-h}^2 \quad (3.47)$$

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

$$\sigma_{+h}^2 = \frac{1}{N(h)} \sum_{i,j \in h} v_j^2 - m_{+h}^2 \quad (3.48)$$

Funcția de variogramă ($\sigma(h)$) 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:

$$\gamma(h) = \frac{1}{2 \cdot N(h)} \sum_{(i,j) \in h} (v_i - v_j)^2 \quad (3.49)$$

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:

$$c(h) = c(-h); \quad \rho(h) = \rho(-h); \quad \gamma(h) = \gamma(-h) \quad (3.50)$$

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 intervariogramă* 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:

$$c_{uv}(h) = \frac{1}{N(h)} \sum_{(i,j) \in h} u_i v_j - m_{u-h} m_{v+h} \quad (3.51)$$

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.

$$m_{u-h} = \frac{1}{N(h)} \sum_{i|_{h_{ij}=h}}^{N(h)} u_i \quad (3.52)$$

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

$$m_{v+h} = \frac{1}{N(h)} \sum_{j|_{h_{ij}=h}}^{N(h)} v_j \quad (3.53)$$

Funcția de intercorelație este dată de ecuația:

$$\rho_{uv}(h) = \frac{c_{uv}(h)}{\sigma_{u-h} \sigma_{v+h}} \quad (3.54)$$

în care:

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

$$\sigma_{u-h}^2 = \frac{1}{N(h)} \sum_{i|_{h_{ij}=h}}^{N(h)} u_i^2 - m_{u-h}^2 \quad (3.55)$$

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

$$\sigma_{v+h}^2 = \frac{1}{N(h)} \sum_{j|_{h_{ij}=h}}^{N(h)} v_j^2 - m_{v+h}^2 \quad (3.56)$$

Funcția de intervareogramă ($\square_{uv}(h)$) se calculează cu relația:

$$\gamma_{uv}(h) = \frac{1}{2N(h)} \sum_{(i,j)|_{h_{ij}=h}}^{N(h)} (u_i - u_j) \cdot (v_i - v_j) \quad (3.57)$$

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 ($c_{uv}(h) \square c_{uv}(-h)$; $\square_{uv}(h) \square \square_{uv}(-h)$), în timp ce funcția de intervareogramă nu depinde de sensul vectorului h ($\square_{uv}(h) = \square_{uv}(-h)$).

1.2.2.3 Analiza staționarității

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

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

$$E\{V(p)\} = m(p) \quad (3.58)$$

$Var\{V(p)\}$ - varianța

$$Var\{V(p)\} = E\{[V(p) - m(p)]^2\} \quad (3.59)$$

$c(p_1, p_2)$ - covarianța

$$c(p_1, p_2) = E\{[V(p_1) - m(p_1)][V(p_2) - m(p_2)]\} \quad (3.60)$$

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

$$2\gamma(p_1, p_2) = Var\{V(p_1) - V(p_2)\} \quad (3.61)$$

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 $(V(p_1), V(p_2))$.

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 :

$$E\{V(p)\} = m, \quad \forall p \quad (3.62)$$

- existența covarianței și invarianța acesteia la translație:

$$c(h) = E\{V(p) \cdot V(p+h)\} - m^2, \quad \forall p \quad (3.63)$$

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

$$Var\{V(p)\} = E\{[V(p) - m]^2\} = c(0), \quad \forall p \quad (3.64)$$

și

$$\gamma(h) = \frac{1}{2} E\{[V(p+h) - v(p)]^2\} = c(0) - c(h), \quad \forall p \quad (3.65)$$

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:

$$\rho(h) = \frac{c(h)}{c(0)} = 1 - \frac{\gamma(h)}{c(0)} \quad (3.66)$$

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:

$$E\{V(p)\} = m, \forall p \quad (3.67)$$

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

$$\text{Var}\{V(p+h) - V(p)\} = E\{[V(p+h) - V(p)]^2\} = 2\gamma(h) \quad (3.68)$$

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:

$$E\{V(p_k)\} \neq E\{V(p_k + h)\}, \quad |h| > r \quad (3.69)$$

Pentru astfel de situații se utilizează *funcțiile structurale* $c(p, p+h)$ sau $\square(p, p+h)$, 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 :

$$c(p_1 - p_2, p_1, p_2) = E\{[V(p_1) - m(p_1)][V(p_2) - m(p_2)]\}, \quad i = 1, \quad j = 2 \quad (3.70)$$

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:

$$V_n = \frac{1}{S_n} \int_{S_n} V(p) dp \quad (3.71)$$

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.

1.2.3 Estimarea punctuală

Estimarea distribuției parametrilor geotehnici a fost realizată prin kriging punctual.

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 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.

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):

$$r_i = v_i^* - v_i, \quad (3.72)$$

Media tuturor erorilor de estimare punctuală este:

$$m_r = \frac{1}{k} \sum_{i=1}^k r_i = \frac{1}{k} \sum_{i=1}^k (v_i^* - v_i). \quad (3.73)$$

Utilizarea expresiei (3.73) 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:

$$V^*(p_0) = \sum_{i=1}^n w_i V(p_i) . \quad (3.74)$$

În mod similar, *erorile de estimare* definite ca diferență între *valoarea estimată* și cea *reală*, ambele *variabile aleatoare*, sunt și ele variabile aleatoare:

$$R(p_0) = V^*(p_0) - V(p_0) . \quad (3.75)$$

Substituind expresia (3.74) în expresia (3.75) se poate exprima eroarea de estimare prin intermediul variabilelor aleatoare originale:

$$R(p_0) = \sum_{i=1}^n w_i V(p_i) - V(p_0) . \quad (3.76)$$

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:

$$E\{R(p_0)\} = E\left\{\sum_{i=1}^n w_i V(p_i) - V(p_0)\right\} = \sum_{i=1}^n w_i \cdot E\{V(p_i)\} - E\{V(p_0)\} = 0 \quad (3.77)$$

Deoarece este presupusă staționaritatea funcției aleatoare, pentru ca estimarea să fie nedeviată suma ponderilor w_i trebuie să fie unitară:

$$E\{R(p_0)\} = \sum_{i=1}^n w_i E\{V\} - E\{V\} = E\{V\} \left(\sum_{i=1}^n w_i - 1 \right) = 0 \quad (3.78)$$

$$\sum_{i=1}^n w_i = 1$$

Varianța erorii de estimare. Varianța erorii de estimare pentru un set de estimări poate fi scrisă sub forma:

$$\sigma_R^2 = \frac{1}{k} \sum_{i=1}^k (r_i - m_r)^2 = \frac{1}{k} \sum_{i=1}^k \left[v_i^* - v_i - \frac{1}{k} \sum_{i=1}^k (v_i^* - v_i) \right]^2 . \quad (3.79)$$

Dacă presupunem și caracterul nedeviat al estimării rezultă că:

$$\sigma_R^2 = \frac{1}{k} \sum_{i=1}^k [v_i^* - v_i]^2 . \quad (3.80)$$

Expresia (3.80) 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_0 unde se realizează estimarea. Estimatorul $V^*(p_0)$ este tot o variabilă aleatoare deoarece este o combinație lineară de variabile aleatoare:

$$V^*(p_0) = \sum_{i=1}^k w_i \cdot V(p_i) . \quad (3.81)$$

Diferența dintre valoarea reală și cea estimată este de asemenea o variabilă aleatoare:

$$R(p_0) = V^*(p_0) - V(p_0) , \quad (3.82)$$

care poate fi dezvoltată sub forma:

$$Var\{R(p_0)\} = Cov\{V^*(p_0) \cdot V^*(p_0)\} - 2Cov\{V^*(p_0) \cdot V(p_0)\} + Cov\{V(p_0) \cdot V(p_0)\} \quad (3.83)$$

Evaluarea rezidului este posibilă doar în cadrul modelului funcției aleatoare staționare. Acest model permite evaluarea covarianței sau variogrammei pentru valorile necunoscute din punctele de estimare pe baza modelului de covarianță sau variogramă dedus din valorile măsurate.

Primul termen din relația (3.83) reprezintă covarianța valorii estimate a variabilei studiate cu ea însăși, adică varianța valorii estimate, ea însăși o combinație lineară de variabile aleatoare:

$$Cov\{V^*(p_0) \cdot V^*(p_0)\} = Var\left\{\sum_{i=1}^n w_i V(p_i)\right\} = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \tilde{c}_{ij} , \quad (3.84)$$

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 (3.83) poate fi descompus sub forma:

$$\begin{aligned}
2Cov\{V^*(p_o) \cdot V(p_o)\} &= 2Cov\left\{\sum_{i=1}^n w_i \cdot V(p_i)\right\} \cdot V(p_o)\} = \\
&= 2E\left\{\sum_{i=1}^n w_i \cdot V(p_i) \cdot V(p_o)\right\} - 2E\left\{\sum_{i=1}^n w_i \cdot V(p_i)\right\} \cdot E\{V(p_o)\} = \\
&= 2\sum_{i=1}^n w_i E\{V(p_i) \cdot V(p_o)\} - 2w_i E\{V(p_i)\} \cdot V(p_o)\} = \\
&= 2\sum_{i=1}^n w_i (E\{V(p_i) \cdot V(p_o)\} - E\{V(p_i)\} \cdot E\{V(p_o)\}) = \\
&= 2\sum_{i=1}^n w_i Cov\{V(p_i) \cdot V(p_o)\} = 2\sum_{i=1}^n w_i \cdot \tilde{c}_{i0}.
\end{aligned} \tag{3.85}$$

Al treilea termen din ecuația (3.83), prin analogie cu primul este varianța valorii reale din punctul de estimare p_o care se exprimă prin intermediul modelului de covarianță sub forma:

$$Cov\{V(p_o) \cdot V(p_o)\} = \tilde{\sigma}^2 \tag{3.86}$$

Pe baza ecuațiilor (3.83), (3.84) și (3.85) se obține expresia varianței erorilor de estimare care permite evaluarea ei pe baza modelului de covarianță :

$$\tilde{\sigma}_R^2 = \tilde{\sigma}^2 + \sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{c}_{ij} - 2\sum_{i=1}^n w_i \cdot \tilde{c}_{i0} , \tag{3.87}$$

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*.

1.2.3.1 Kriging punctual ordinar

Calculul ponderilor din combinația liniară 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*.

Ecuațiile 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:

$$m_R = 0 \tag{3.88}$$

și a minimului varianței erorii de estimare:

$$\sigma_R^2 - \text{minimum} \quad (3.89)$$

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 (3.88). Asigurarea condiționării suplimentare din ecuația (3.89) 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 (3.87) cu condiția de estimare nedeviată din ecuația (3.88) 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 (3.88) numită parametrul lui Lagrange (μ):

$$\tilde{\sigma}_R^2 = \tilde{\sigma}^2 + \sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{c}_{ij} - 2 \sum_{i=1}^n w_i \cdot \tilde{c}_{i0} + 2\mu \left(\sum_{i=1}^n w_i - 1 \right). \quad (3.90)$$

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.

Ecuația varianței erorii de estimare în forma (3.90) 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:

$$\begin{cases} \frac{\partial(\tilde{\sigma}_R^2)}{\partial w_1} = 0 \\ \vdots \\ \frac{\partial(\tilde{\sigma}_R^2)}{\partial w_n} = 0 \\ \frac{\partial(\tilde{\sigma}_R^2)}{\partial \mu} = 0 \end{cases} \quad (3.91)$$

Calculul derivatei în raport cu w_1 , desfășurat separat pentru cei patru termeni ai varianței erorii de estimare dată de ecuația (3.90), conduce la următoarele rezultate:

- primul termen:

$$\frac{\partial(\tilde{\sigma}^2)}{\partial w_1} = 0 ; \quad (3.92)$$

- al doilea termen:

$$\frac{\partial \left(\sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{c}_{ij} \right)}{\partial w_1} = \frac{\partial \left(w_1^2 \tilde{c}_{11} + 2w_1 \sum_{j=2}^n w_j \cdot \tilde{c}_{1j} \right)}{\partial w_1} = ; \quad (3.93)$$

$$= 2w_1 \tilde{c}_{11} + 2 \sum_{j=2}^n w_j \cdot \tilde{c}_{1j} = 2 \sum_{j=1}^n w_j \cdot \tilde{c}_{1j}$$

- al treilea termen:

$$\frac{\partial \left(2 \sum_{i=1}^n w_i \cdot \tilde{c}_{i0} \right)}{\partial w_1} = 2 \tilde{c}_{10} ; \quad (3.94)$$

- al patrulea termen:

$$\frac{\partial \left(2\mu \left(\sum_{i=1}^n w_i - 1 \right) \right)}{\partial w_1} = 2\mu . \quad (3.95)$$

Prin combinarea ecuațiilor (3.93-3.95) se obține expresia pentru derivata în raport cu w_1 :

$$\frac{\partial (\tilde{\sigma}_R^2)}{\partial w_1} = 2 \sum_{j=1}^n w_j \cdot \tilde{c}_{1j} - 2 \tilde{c}_{10} + 2\mu . \quad (3.96)$$

Primele n derivate în raport cu w_i ($i = 1, 2, \dots, n$) au forma generală:

$$\sum_{j=1}^n w_j \cdot \tilde{c}_{ij} + \mu = \tilde{c}_{i0}, \forall i = 1, 2, \dots, n . \quad (3.97)$$

Derivata în raport cu parametrul lui Lagrange, a $(n+1)$ -a derivată, are forma:

$$\frac{\partial (\tilde{\sigma}_R^2)}{\partial \mu} = \frac{\partial \left(2\mu \left(\sum_{i=1}^n w_i - 1 \right) \right)}{\partial \mu} = 2 \left(\sum_{i=1}^n w_i - 1 \right) . \quad (3.98)$$

Sistemul de kriging ordinar se obține prin anularea celor $(n+1)$ derivate parțiale date de ecuațiile (3.97) și (3.98) și are forma:

$$\left\{ \begin{array}{l} \frac{\partial (\tilde{\sigma}_R^2)}{\partial w_1} = 2 \sum_{i=1}^n w_i \cdot \tilde{c}_{1i} - 2 \tilde{c}_{10} + 2\mu = 0 \\ \vdots \\ \frac{\partial (\tilde{\sigma}_R^2)}{\partial w_n} = 2 \sum_{i=1}^n w_i \cdot \tilde{c}_{ni} - 2 \tilde{c}_{n0} + 2\mu = 0 \\ \frac{\partial (\tilde{\sigma}_R^2)}{\partial \mu} = 2 \sum_{i=1}^n w_i - 1 = 0 \end{array} \right. \quad (3.99)$$

Prin separarea coeficienților cunoscuți și a necunoscutelor, sistemul de kriging poate fi scris sub forma matricială:

$$\begin{bmatrix} \tilde{c}_{11} & \tilde{c}_{12} & \cdots & \tilde{c}_{1n} & 1 \\ \tilde{c}_{21} & \tilde{c}_{22} & \cdots & \tilde{c}_{2n} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \tilde{c}_{n1} & \tilde{c}_{n2} & \cdots & \tilde{c}_{nn} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} \tilde{c}_{i_0} \\ \tilde{c}_{j_0} \\ \vdots \\ \tilde{c}_{n_0} \\ 1 \end{bmatrix} . \quad (3.100)$$

Dacă notăm cu C matricea coeficienților, cu W vectorul necunoscutelor și cu D vectorul termenilor liberi, sistemul (3.100) poate fi scris sub forma:

$$C \cdot W = D \quad (3.101)$$

a cărei soluție este:

$$W = C^{-1} \cdot D \quad (3.102)$$

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 (3.90), dar pentru a găsi o expresie în raport numai cu valorile măsurate se pleacă de la ecuația (3.97) în care ambii membri se multiplică cu w_i :

$$w_i \left(\sum_{j=1}^n w_j \cdot \tilde{c}_{ij} + \mu \right) = w_i \cdot \tilde{c}_{i_0} \quad (3.103)$$

și se însumează pentru toate cele n puncte de observație, rezultând:

$$\sum_{i=1}^n w_i \sum_{j=1}^n w_j \cdot \tilde{c}_{ij} + \sum_{i=1}^n w_i \cdot \mu = \sum_{i=1}^n w_i \cdot \tilde{c}_{i_0} , \quad (3.104)$$

care sub forma:

$$\sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{c}_{ij} = \sum_{i=1}^n w_i \cdot \tilde{c}_{i_0} + \mu \quad (3.105)$$

se înlocuiește în (3.90) și se obține:

$$\tilde{\sigma}_R^2 = \tilde{\sigma}^2 - \left(\sum_{i=1}^n w_i \cdot \tilde{c}_{i_0} + \mu \right) , \quad (3.106)$$

care în formă matricială este:

$$\tilde{\sigma}_R^2 = \tilde{\sigma}^2 - w \cdot D . \quad (3.107)$$

Ecuatiile sistemului de kriging în funcție de variogramă. În aceeași ipoteze care au permis deducerea expresiei varianței erorii de estimare (3.90) poate fi utilizată și *variograma* a cărei relație de definiție este:

$$\gamma_{ij} = \frac{1}{2} E\{(V(p_i) - V(p_j))^2\}, \quad (3.108)$$

care pentru evaluarea varianței erorii de estimare poate fi scrisă sub forma:

$$\begin{aligned} \gamma_{ij} &= \frac{1}{2} E\{[(V(p_i) - V(p_0)) - (V(p_j) - V(p_0))]^2\} = \\ &= \frac{1}{2} E\{[V(p_i) - V(p_0)]^2\} + \frac{1}{2} E\{[V(p_j) - V(p_0)]^2\} - \\ &\quad - E\{[V(p_i) - V(p_0)][V(p_j) - V(p_0)]\} = \\ &= \tilde{\gamma}_{i0} + \tilde{\gamma}_{j0} - E\{[V(p_i) - V(p_0)][V(p_j) - V(p_0)]\} \end{aligned} \quad (3.109)$$

Varianța erorii de estimare în această variantă este dată de relația:

$$\begin{aligned} \tilde{\sigma}_R^2 &= E\{[V^*(p_0) - V(p_0)]^2\} = E\left\{\left[\sum_{i=1}^n w_i \cdot V(p_i) - \sum_{i=1}^n w_i \cdot V(p_0)\right]^2\right\} = \\ &= \sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot E\{[V(p_i) - V(p_0)][V(p_j) - V(p_0)]\} \end{aligned} \quad (3.110)$$

în care, utilizând forma variogramei din ecuația (3.109), se obține:

$$\begin{aligned} \tilde{\sigma}_R^2 &= E\{[V^*(p_0) - V(p_0)]^2\} = -\sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{\gamma}_{ij} + \\ &\quad + \sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{\gamma}_{i0} + \sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{\gamma}_{j0} = \\ &= -\sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{\gamma}_{ij} + \sum_{j=1}^n w_j \left(\sum_{i=1}^n w_i \cdot \tilde{\gamma}_{i0} + \sum_{i=1}^n w_i \cdot \tilde{\gamma}_{j0} \right) = \\ &= -\sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{\gamma}_{ij} + 2 \sum_{i=1}^n w_i \cdot \tilde{\gamma}_{i0} \end{aligned} \quad (3.111)$$

Pentru minimizarea varianței erorii de estimare, în condiția de estimare nedeviată (3.88), utilizând parametrul lui Lagrange (μ), relația corespunzătoare ecuației (3.90) scrisă pentru variogramă este:

$$\tilde{\sigma}_R^2 = -\frac{1}{2} \left(\sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{\gamma}_{ij} + 2 \sum_{i=1}^n w_i \cdot \tilde{\gamma}_{i0} \right) - \mu \left(\sum_{i=1}^n w_i - 1 \right) \quad (3.112)$$

Minimizarea varianței erorii de estimare scrisă sub forma (3.112) 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 (3.93-3.96) se obțin derivatele în raport cu ponderile w sub forma:

$$\frac{\partial(\tilde{\sigma}_R^2)}{\partial w_i} = \frac{1}{2} \left(-2 \sum_{j=1}^n w_j \cdot \tilde{\gamma}_{ij} + 2\tilde{\gamma}_{i0} \right) - \mu, \forall i = 1, 2, \dots, n \quad (3.113)$$

iar pentru derivata în raport cu μ

$$\frac{\partial(\tilde{\sigma}_R^2)}{\partial \mu} = -\sum_{i=1}^n w_i + 1 \quad (3.114)$$

Sistemul de kriging în raport cu variograma se obține prin anularea celor $(n+1)$ derivate parțiale din ecuațiile (3.113) și (3.114) și are forma:

$$\begin{cases} \sum_{j=1}^n w_j \cdot \tilde{\gamma}_{ij} + \mu = \tilde{\gamma}_{i0}, \forall i = 1, 2, \dots, n \\ \sum_{i=1}^n w_i = 1 \end{cases} \quad (3.115)$$

care sub formă matricială poate fi scris:

$$\begin{bmatrix} \tilde{\gamma}_{11} & \tilde{\gamma}_{12} & \dots & \tilde{\gamma}_{1n} & 1 \\ \tilde{\gamma}_{21} & \tilde{\gamma}_{22} & \dots & \tilde{\gamma}_{2n} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \tilde{\gamma}_{n1} & \tilde{\gamma}_{n2} & \dots & \tilde{\gamma}_{nn} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} \tilde{\gamma}_{10} \\ \tilde{\gamma}_{20} \\ \vdots \\ \tilde{\gamma}_{n0} \\ 1 \end{bmatrix} \quad (3.116)$$

Valoarea minimă a *varianței erorii de estimare* în raport cu valorile măsurate este dată de relația:

$$\tilde{\sigma}_R^2 = \sum_{i=1}^n w_i \cdot \tilde{\gamma}_{i0} + \mu \quad (3.117)$$

dedusă în mod analog cu relația (3.106).

Ecuțiile sistemului de kriging în funcție de corelogramă. Între corelogramă și covarianță există relația:

$$\tilde{\rho}_{ij} = \frac{\tilde{c}_{ij}}{\tilde{\sigma}^2}, \quad (3.118)$$

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 (3.99) și (3.116) sistemul poate fi scris:

$$\begin{cases} \sum_{j=1}^n w_j \cdot \tilde{\rho}_{ij} + \mu = \tilde{\rho}_{i0}, \forall i=1,2,\dots,n \\ \sum_{i=1}^n w_i = 1 \end{cases} \quad (3.119)$$

sau sub formă matricială:

$$\begin{bmatrix} \tilde{\rho}_{11} & \tilde{\rho}_{12} & \cdots & \tilde{\rho}_{1n} & 1 \\ \tilde{\rho}_{21} & \tilde{\rho}_{22} & \cdots & \tilde{\rho}_{2n} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \tilde{\rho}_{n1} & \tilde{\rho}_{n2} & \cdots & \tilde{\rho}_{nn} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} \tilde{\rho}_{10} \\ \tilde{\rho}_{20} \\ \vdots \\ \tilde{\rho}_{n0} \\ 1 \end{bmatrix}. \quad (3.120)$$

În cazul utilizării corelogramei *varianța minimă a erorii de estimare* se calculează cu relația:

$$\tilde{\sigma}_R^2 = \sum_{i=1}^n w_{i0} \cdot \tilde{\rho}_{i0} + \mu \quad (3.121)$$

În practică, datorită flexibilității variogramei se preferă utilizarea acesteia pentru estimarea *ponderilor* (w_i) (sistemul **3.116**) și a *varianței minime de estimare* ($\tilde{\sigma}_k^2$) (ecuația **3.117**).

1.2.3.2 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:

$$E\{V(p)\} = m(p) \quad (3.122)$$

Funcția aleatoare poate fi descompusă într-o tendință ($m(p)$) și un termen rezidual $Y(p)$ staționar sau nestaționar dar cu speranța matematică nulă:

$$V(p) = m(p) + Y(p) \text{ cu } E\{Y(p)\} = 0 \quad (3.123)$$

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.3.8**).

Forma tendinței este o combinație lineară de K funcții cunoscute cu coeficienți necunoscuți de ecuație:

$$m(p) = \sum_{i=1}^K a_i \cdot f_i(p) \quad (3.124)$$

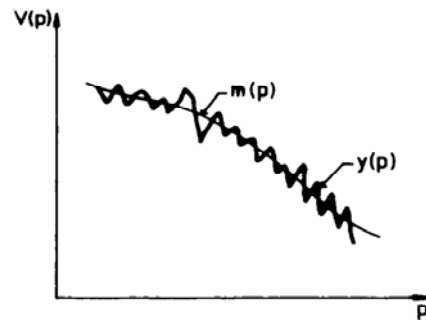


Fig. 3.8. Modelul funcției aleatoare nestaționare

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*):

$$m(p) = a_1 + a_2 \cdot p \quad (3.125)$$

sau doi (*derivă pătratică*):

$$m(p) = a_1 + a_2 \cdot p + a_3 \cdot p^2 \quad (3.126)$$

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:

$$\begin{cases} m(t) = a_1 + a_2 \cdot t \\ m(t) = a_1 + a_2 \cdot t + a_3 \cdot t^2 \end{cases} \quad (3.127)$$

- dacă spațiul este *bidimensional*, reprezentat într-un sistem de referință de coordonate rectangulare x și y , formele celor două tendințe sunt:

$$\begin{cases} m(x, y) = a_1 + a_2 \cdot x + a_3 \cdot y \\ m(x, y) = a_1 + a_2 \cdot x + a_3 \cdot y + a_4 \cdot x^2 + a_5 \cdot x \cdot y + a_6 \cdot y^2 \end{cases} \quad (3.128)$$

- dacă spațiul este *tridimensional*, de coordonate x , y și z forma tendințelor este:

$$\begin{cases} m(x, y, z) = a_1 + a_2 \cdot x + a_3 \cdot y + a_4 \cdot z \\ m(x, y, z) = a_1 + a_2 \cdot x + a_3 \cdot y + a_4 \cdot z + a_5 \cdot x^2 + a_6 \cdot y^2 + a_7 \cdot z^2 + a_8 \cdot x \cdot y + a_9 \cdot x \cdot z + a_{10} \cdot y \cdot z \end{cases} \quad (3.129)$$

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:

$$m(x, y, z) = a_1 + a_2 \cdot z + a_3 \cdot z^2 \quad (3.130)$$

Covarianța și variograma funcției aleatoare nestacionare. Pentru funcția aleatoare nestacionară cu structura din ecuația (3.123) formulele covarianței și variogramei sunt:

$$C(p_1, p_2) = E\{V(p_1) \cdot V(p_2)\} - m(p_1) \cdot m(p_2) = E\{Y(p_1) \cdot Y(p_2)\} \quad (3.131)$$

și:

$$\begin{aligned} 2\gamma(p_1, p_2) &= Var\{V(p_1) - V(p_2)\} = E\{[V(p_1) - V(p_2)]^2\} - \\ &- [m(p_1) - m(p_2)]^2 = Var\{Y(p_1) - Y(p_2)\} = E\{[Y(p_1) - Y(p_2)]^2\} \end{aligned} \quad (3.132)$$

Variograma din formula (2.81), adică aceea a reziduuului *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:

$$V_0^* = \sum_{i=1}^n w_i \cdot v_i. \quad (3.133)$$

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 (3.124) condiția de estimare nedeviată devine:

$$\sum_{i=1}^n w_i \sum_{l=1}^k a_l \cdot f_l(p_i) - \sum_{l=1}^k a_l \cdot f_l(p_0) = 0, \quad (3.134)$$

din care, deoarece coeficienții derivatei sunt necunoscuți, trebuie ca:

$$\sum_{i=1}^n w_i \cdot f_l(p_i) = f_l(p_0), \quad (3.135)$$

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:

$$\tilde{\sigma}_R^2 = 2 \sum_{i=1}^n w_i \cdot \tilde{\gamma}_{i0} - \sum_{i=1}^n \sum_{j=1}^n w_i \cdot w_j \cdot \tilde{\gamma}_{ij} \quad (3.136)$$

astfel încât pentru minimizarea ei în condițiile unei estimări nedeviate *sistemul de kriging universal* este:

$$\left\{ \begin{array}{l} \sum_{j=1}^n w_j \cdot \tilde{\gamma}_{ij} + \sum_{l=1}^k \mu_l \cdot f_l(p_i) = \gamma_{i0}, i = 1, 2, \dots, n \\ \sum_{i=1}^n w_i \cdot f_l(p_i) = f_l(p_0), l = 1, 2, \dots, k \end{array} \right. , \quad (3.137)$$

sau sub formă matricială, în cazul unei tendințe de forma (3.128) pentru un spațiu bidimensional:

$$\begin{bmatrix} \tilde{\gamma}_{11} & \tilde{\gamma}_{12} & \cdots & \tilde{\gamma}_{1n} & 1 & x_1 & y_1 \\ \tilde{\gamma}_{21} & \tilde{\gamma}_{22} & \cdots & \tilde{\gamma}_{2n} & 1 & x_2 & y_2 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ \tilde{\gamma}_{n1} & \tilde{\gamma}_{n2} & \cdots & \tilde{\gamma}_{nn} & 1 & x_n & y_n \\ 1 & 1 & \cdots & 1 & 0 & 0 & 0 \\ x_1 & x_2 & \cdots & x_n & 0 & 0 & 0 \\ y_1 & y_2 & \cdots & y_n & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} \tilde{\gamma}_{10} \\ \tilde{\gamma}_{20} \\ \vdots \\ \tilde{\gamma}_{n0} \\ 1 \\ x_0 \\ y_0 \end{bmatrix} \quad (3.138)$$

Relația de calcul pentru varianța erorii de estimare este:

$$\tilde{\sigma}_R^2 = \sum_{i=1}^n w_i \cdot \tilde{\gamma}_{i0} + \sum_{j=1}^k \mu_j \cdot f_j(p_0). \quad (3.139)$$

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.