

MEODICA CERCETĂRII

1. METODICA CERCETĂRII	2
1.1. Entropia rețelelor de investigare	3
1.1.1. Entropia 2D a rețelelor de monitorizare.....	4
1.1.1.1. Noțiuni elementare	4
1.1.1.2. Aplicație 2D.....	6
1.1.1.3. Concluzii 2D	7
1.1.2. Entropia 1D a rețelelor de monitorizare.....	8
1.1.2.1. Noțiuni elementare	9
1.1.2.2. Aplicație 1D.....	11
1.1.2.3. Concluzii 1D	13
1.2. Metoda punctului fictiv.....	14
1.2.1. Noțiuni elementare.....	14
1.2.2. Etapele de aplicare ale metodei punctului fictiv	14
1.2.3. Aplicație 2D	15
BIBLIOGRAFIE	20

1. METODICA CERCETĂRII

Metodica cercetării pe lângă metodele specifice de cercetare a factorilor proceselor geologice are o logistică unitară pentru doate domeniile de cercetare ale Științelor Pământului:

- Stabilirea **obiectivului** cercetării procesului geologic
- Identificarea factorilor care condiționează desfășurarea procesului, factori care configurează modelul conceptual al procesului cu cele trei componente:
 - **Spațiul și intervalul de timp** în care se desfășoară procesul geologic
 - **Caracteristicile parametrice** ale spațiului în care se manifestă procesul geologic
 - **Energia** care susține desfășurarea procesului geologic
- **Modelarea corelațiilor** dintre obiectul cercetării și factorii care îl condiționează

Cercetarea se bazează pe o **rețea de investigare** a cărei configurare are ca obiectiv surprinderea cu erori minime a variabilității spațiale și temporale a variabilelor regionalizate ale proceselor geologice cu:

- număr minim de puncte observație
- frecvență minimă a măsurătorilor în punctele de investigare

Configurarea rețelei este în funcție de numărul variabilelor regionalizate și de eroarea admisibilă pentru evaluarea distribuției spațiale a tuturor variabilelor regionalizate monitorizate.

Datele minime necesare configurării unei rețele de investigare pe baza unei investigații preliminare sunt:

- **coordonatele** punctelor de observație inițiale (N) ale rețelei de investigare ($i=1,2,\dots,N$):
 - (X_i, Y_i)
- **valorile variabilei/variabilelor** monitorizate la momentul inițial în toate punctele de observație:
 - $V(T_o)$
- **serii de timp** de valori în fiecare punct de observație:
 - $V(X_1, Y_1, T_1), V(X_1, Y_1, T_2), \dots, V(X_1, Y_1, T_{50})$
 -
 - $V(X_N, Y_N, T_1), V(X_N, Y_N, T_2), \dots, V(X_N, Y_N, T_{50})$
- **abaterea standard maximă acceptată (KSD)** (KSD : Kriging Standard Deviation) impusă de eroarea maximă admisă de evaluare a distribuției spațiale și temporale în zona investigată, corespunzător unui risc asumat.

1.1. Entropia rețelelor de investigare

Utilizarea entropiei Shannon (**Fig.1**) în proiectarea rețelelor de monitorizare are ca obiective:

- evaluarea **incertitudinii medii 2D** privind variabilitatea în spațiu a **variabilelor aleatoare regionalizate**, etapa preliminară aplicării metodei topo-probabiliste a **punctului fictiv** (D.Scrădeanu et.al., 2001, 2003), pentru ameliorarea eficienței rețelelor de monitorizare
- evaluarea **entropiei 1D a lanțurilor Markov**, utilizată la stabilirea **intervalului de eșantionare al seriilor de timp** pentru variabilele monitorizate.

Metodologia de calcul și aplicațiile de la paragrafele **§1.2** și **§2.2** sunt completate cu două fișiere de tip excel postate pe site-ul dedicat proiectării rețelelor de monitorizare:

1. Fișierul **ENTROPIE_2D.xls** permite calculul entropiei_2D pentru o serie de 55 de valori care trebuie plasate pe **ZONA VERDE** de pe coloana V(To) după o procedură similară, descrisă detaliat pentru entropia 1D
2. Fișierul **ENTROPIE_1D.xls** permite calculul entropiei_1D pentru o serie de 50 de valori separate în trei grupe valorice (A,B,C), după următoarea procedură:
 - a. se deschide fișierul **ENTROPIE_1D.xls** care aplică metodologia pentru o serie de 50 de valori plasata pe coloana V(Ti) în **ZONA VERDE**
 - b. se șterg valorile seriei test
 - c. se plasează pe coloana **ZONA VERDE** seria de valori pentru care se dorește calculul entropiei 1D

3. NOTA:

- a. Cele trei grupe valorice (A,B,C) sunt stabilite prin divizarea amplitudinii de variație a variabilei studiate în trei interval egale (se pot alege și alte variante de lucru prin modificarea corespunzătoare a limitelor intervalelor celor trei stări: A, B, C) .
- b. Algoritmul de calcul este explicat în paragrafele **§1.2** și **§2.2**
- c. Dacă seria de valori este mai mare de 50/55 de valori trebuie extinse formulele de calcul de pe **coloanele MNOPQRS/ M,N,...,AB**, până la rândul corespunzător ultimei valori a seriei de valori prelucrate.

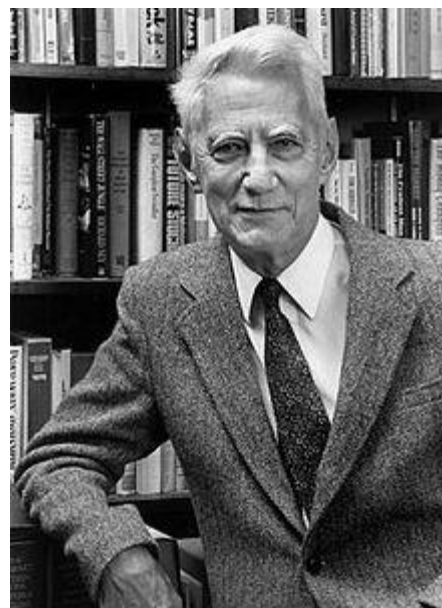


Fig.1. Claude Shannon, părintele teoriei informației

1.1.1. Entropia 2D a rețelelor de monitorizare

1.1.1.1. **Noțiuni elementare**

Entropia lui Shannon, inspirată din termodinamică, în context **probabilistic**, este o măsură a informației conținută în distribuția **variabilei regionalizate** (V), ce urmează să fie monitorizată în stațiile rețelei de monitorizare ($i = 1, 2, \dots, N$; **Fig.1.1**):

$$V : \begin{pmatrix} V_1 & V_2 & \dots & V_N \\ p_1 & p_2 & \dots & p_N \end{pmatrix}; \quad H(V) = - \sum_{i=1}^{i=N} p_i \cdot \log_2 p_i \quad (1.1)$$

p_i este probabilitatea de realizare a valorii din stația V_i

Pentru calculul entropiei se utilizează **logaritmul** pentru a permite **însurarea** incertitudinilor unor variabile independente (U, V): $H(U, V) = H(U) + H(V)$

Unitatea de măsură în care se exprimă entropia este **shannon-ul/ bit-ul**, dacă baza în care se calculează logaritmul este 2. Conținutul informațional al unui eveniment cu probabilitatea $p = \frac{1}{2}$ este 1 shannon:

$$1sh = - \sum_{i=1}^{i=2} \frac{1}{2} \cdot \log_2 \frac{1}{2} = 1$$

Pentru exemplificare modului de calcul al entropiei Shannon ($H(V)$) în context probabilistic, aplicat unei rețele de monitorizare, considerăm **două situații extreme (Fig.1.1)** în **N=12 puncte de observație** ale unei rețele de monitorizare care identifică:

- a) **N=1 stare/valoare** posibilă a variabilei V : **A (Fig.1.1.a)**
- b) **N=12 stări/valori** egal posibile ale variabilei V : **A,B,C,D,E,F,G,H,I,J,K,L (Fig.1.1.a)**.

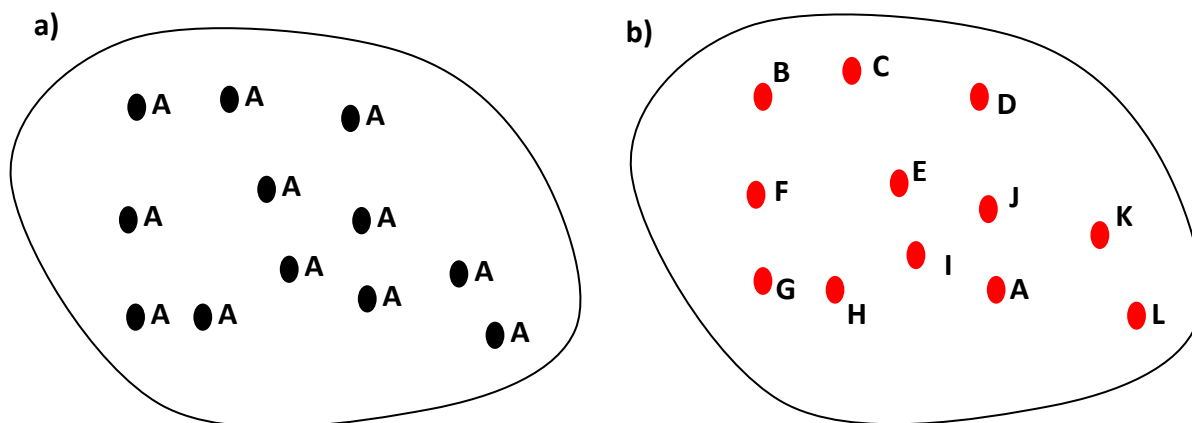


Fig. 1.1. Două situații extreme pentru exemplificarea modului de calcul al entropiei lui Shannon în context probabilistic, aplicat rețelelor de monitorizare

Tablourile variabilei regionalizate (V) din cele 12 puncte de observație, **la un moment dat** (t_j), corespunzător celor două situații extreme sunt:

- a) **cunoasterea completă** a variabilei, cand **o singură stare este prezentă** (starea A) cu probabilitatea unitară (repartiție uniformă):

$$V: \begin{pmatrix} A & A & A & A & A & A & A & A & A & A & A & A \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}; p_A = \frac{12}{12} = 1$$

situație în care entropia este MINIMĂ și egală cu ZERO:

$$H(V) = -\sum_{i=1}^{i=12} p_i \cdot \log_2 p_i = -\sum_{i=1}^{i=12} 1 \cdot \log_2(1) = 0$$

- b) **incertitudine maximă** asupra stării variabilei, cand **toate cele 12 stări posibile** (A,B,C,D,E,F,G,H,I,J,K,L) au probabilități egale:

$$V: \begin{pmatrix} A & B & C & D & E & F & G & H & I & J & K & L \\ \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} \end{pmatrix}; p_A = p_B = p_C = \dots p_L = \frac{1}{12}$$

situație în care entropia este MAXIMĂ și egală cu $\log_2 12$:

$$H(V) = -\sum_{i=1}^{i=12} p_i \cdot \log_2 p_i = -\sum_{i=1}^{i=12} \frac{1}{12} \cdot \log_2 \left(\frac{1}{12}\right) = \log_2 12 = 3.58$$

Variația entropiei între cele două extreme, este similară cu variația funcției $y = x \cdot \log x$.

Pentru calcule (în excel), conform graficului (**Fig.1.2**), vom considera $0 \cdot \log_2 0 = 0$ (N – numărul de stări/valori distincte al variabilei V ; detalii Keth Konrad, *Probability distribution and maximum entropy*, [4]).

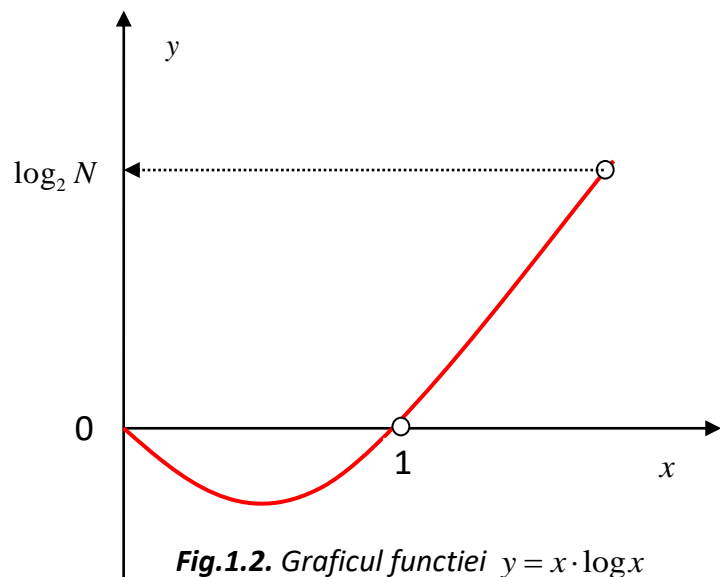


Fig.1.2. Graficul funcției $y = x \cdot \log x$ (dupa K.Conrad cu completări)

1.1.1.2. Aplicație 2D

Să se calculeze **entropia 2D** a unei rețele de monitorizare formată din **55 de puncte de observație**, pe baza valorilor ($V(T_0)$); **Tabelul 1.1**).

Tabelul 1.1. Valorile variabilei V la momentul T_0 în cele 55 de puncte de monitorizare

NR	X	Y	V(T_0)	NR	X	Y	V(T_0)
1	79.23	535.13	24.85	29	862.42	41.88	68.30
2	125.20	577.56	30.38	30	867.72	149.72	73.61
3	54.48	570.49	27.34	31	973.80	20.66	86.13
4	43.87	480.32	19.35	32	873.03	317.67	85.17
5	134.04	404.30	16.41	33	770.49	365.41	72.73
6	171.16	538.66	28.73	34	706.84	421.98	68.34
7	96.91	434.36	17.29	35	625.52	453.80	60.38
8	40.34	312.37	8.39	36	530.05	489.16	52.37
9	123.43	303.53	9.93	37	473.48	551.04	52.46
10	58.02	365.41	11.69	38	650.27	565.18	74.83
11	174.70	215.13	7.64	39	657.34	499.77	68.91
12	54.48	172.70	2.97	40	706.84	545.74	80.57
13	248.95	77.24	6.68	41	756.34	566.95	90.28
14	252.49	255.80	12.99	42	966.73	178.01	91.98
15	319.67	114.36	11.40	43	970.26	338.89	103.98
16	266.63	317.67	17.11	44	922.53	370.71	98.05
17	400.99	41.88	15.18	45	970.26	271.71	98.73
18	132.27	34.81	1.83	46	835.90	598.77	106.81
19	522.98	33.04	25.30	47	869.49	572.25	109.40
20	565.41	98.45	31.32	48	906.62	593.47	118.53
21	662.64	10.06	39.82	49	982.64	529.82	125.22
22	660.88	75.47	41.37	50	966.73	490.93	117.86
23	423.97	475.02	40.42	51	940.21	467.95	110.50
24	56.25	36.57	0.46	52	869.49	444.96	95.69
25	22.66	110.83	1.12	53	957.89	406.07	107.65
26	512.37	381.32	41.26	54	956.12	554.58	123.03
27	0.00	0.00	0.00	55	1000.00	600.00	137.05
28	765.18	71.93	54.88				

Rezolvare 2D:(ENTROPIE_2D.xls)

Pentru calculul **entropiei 2D** se definesc trei grupe valorice, divizând amplitudinea de variație în trei intervale egale (**Fig.1.3**):

- $A \in [0;45,69)$
- $B \in [45,69;91,37)$
- $C \in [91,37;137,06]$

Numărul grupelor valorice se definește în funcție de gradul de detaliu la care se dorește cunoașterea variației medii globale a variabilei în suprafața monitorizată.

Compararea entropiei aceleiași variabile pentru un număr diferit de grupe valorice este utilizată pentru a sesiza continuitatea medie în spațiul monitorizat.

Se pot face zonări spațiale ale suprafeței monitorizate pentru a sesiza zone de incertitudini diferite, pentru același număr de grupe și pentru aceeași variabilă.

Entropia variabilei $V(T0)$, calculată pentru cele trei grupe (**Fig.1.3**), utilizând formula lui Shannon (1.1), este:

$$H(V) = -\sum_{i=1}^{i=3} p_i \cdot \log_2 p_i = -\left[\frac{27}{55} \cdot \log_2 \left(\frac{27}{55} \right) + \frac{14}{55} \cdot \log_2 \left(\frac{14}{55} \right) + \frac{14}{55} \cdot \log_2 \left(\frac{14}{55} \right) \right] = 1,51$$

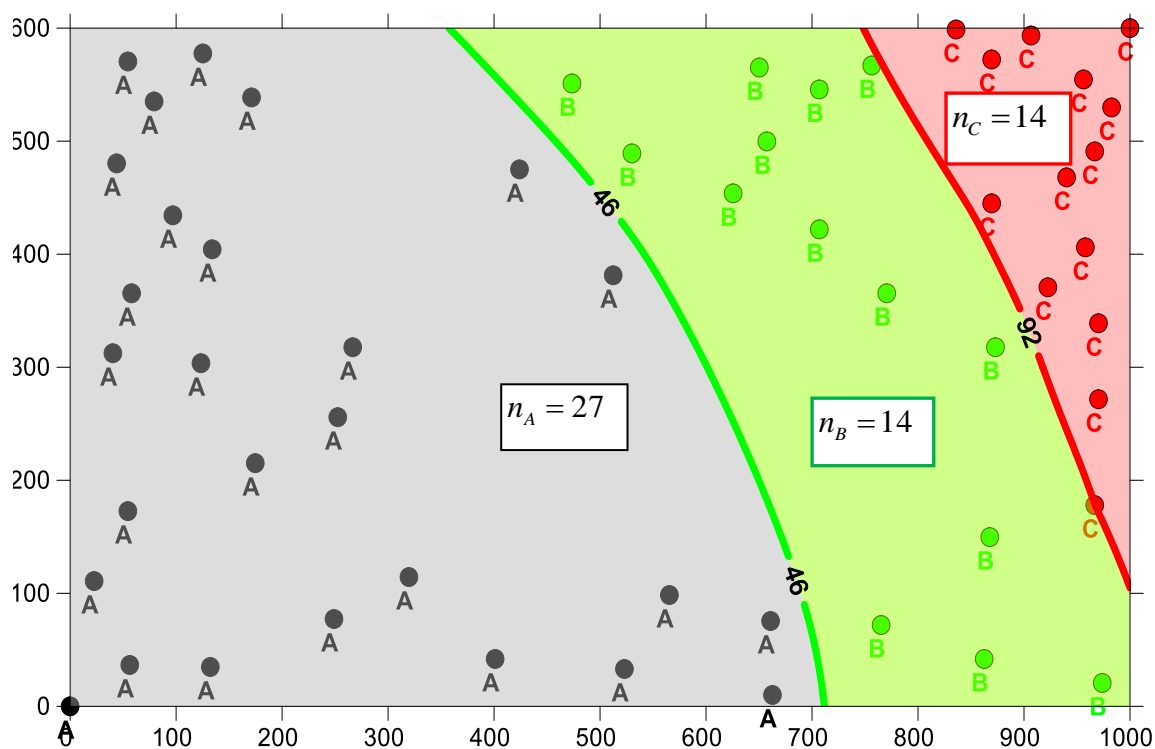


Fig.1.3. Poziția celor 55 de puncte de observație din grupele A,B și C

1.1.1.3. Concluzii 2D

Valoarea “mare” a entropiei rețelei de monitorizare (relativ la $H_{MAX} = \log_2 3 = 1,59$) indică o INCERTITUDINE RIDICATĂ asupra variabilității spațiale a variabilei $V(T0)$ recomandând completarea rețelei cu **puncte suplimentare de observație**.

Completarea eficientă a rețelei de monitorizare se poate face prin **metoda punctului fictiv**.

1.1.2. Entropia 1D a rețelelor de monitorizare

Variația în timp a valorii unei variabile într-o stație a rețelei de monitorizare este rezultatul interferenței unui număr mare de factori, număr de factori cu atât mai mare cu cât variabila măsurată reflectă un proces mai complex.

Cota nivelului piezometric al unui acvifer freatic, spre exemplu, este rezultatul interacțiunii unui număr mare de factori (curgerea apelor subterane fiind un proces complex):

- precipitațiile din zona de alimentare a acviferului
- temperatura aerului
- acoperirea vegetală/tipul de utilizare a terenului
- panta terenului
- constituția litologică a formațiunilor din zona vadoasă
- grosimea depozitelor permeabile
- umiditatea zonei vadoase
- conductivitatea hidraulică a acviferului
- etc.

Este oneroasă tentativa construirii unui **model funcțional** care să permită simularea variației cotei nivelului piezometric al acviferului **în funcție de toți factorii** care o condicionează.

Pentru studiul variației în timp a variabilelor care sunt rezultatul unui proces complex se apelează la **modele statistice** în care se modelează variația în **timp** a valorilor **variabilei rezultative, variabilele factoriale** fiind închise într-o “**cutie neagră**”.

Lanturile Markov sunt un model statistic adecvat analizei seriilor de timp complexe, model care presupune evaluarea a două mărimi:

- **probabilitatea** de trecere dintr-o stare în alta
- **entropia** asociată fiecărei schimbări de stare.

Evaluarea **entropiei lanturilor Markov** este utilizată la alegerea **intervalului de eșantionare al seriilor de timp** pentru variabilele monitorizate.

Intervalul de eșantionare (Δt) al seriilor de timp pentru variabilele monitorizate este:

- **invers proporțional** cu valoarea **entropiei** matricii probabilităților de tranziție
- **direct proporțional** cu **eroarea admisibilă** $\varepsilon(\alpha)$ la un **risc asumat** (α).

1.1.2.1. Noțiuni elementare

Un **process Markov/lanț Markov (Fig.2.1)** este un process stochastic în care starea/valoarea curentă reține toată informația despre întreaga evoluție a procesului.

Procesul **stochastic/aleatoriu**, este un model care cuantifică incertitudinea evoluției în timp a unui proces natural complex, condiționat de *un mare număr de variabile factoriale*. Aceasta înseamnă că deși se cunoaște starea inițială a unui proces, există mai multe posibilități de continuare a procesului, dar unele căi sunt mai probabile decât altele.

Matricea de tranziție a unui proces Markov este instrumentul operațional care permite identificarea componentei “corelaționale” a procesului stochastic (Daniel Scradeanu, 1995, Informatica geologică, Ed.Univ.Buc.).

Un process Markov având ca rezultat doar două valori/stări distincte (**A,B**: $N_s = 2$) ale variabilei principale, descris de o succesiune de 10 valori (**Fig.2.2**) este caracterizat prin:

- Numarul total de valori disponibile(**A+B**):
 - $N = 10$
- Frecvența absolută a celor două stări (**A** si **B**: $N_s = 2$):
 - **A**: $n_A = 6$
 - **B**: $n_B = 4$
- Probabilitatea medie pentru fiecare stare:
 - **A**: $p_A = \frac{n_A}{N} = \frac{6}{10}$
 - **B**: $p_B = \frac{n_B}{N} = \frac{4}{10}$
- Tipurile de tranziții:
 - **Tranziție A->A**
 - **Tranziție:A->B**
 - **Tranziție:B->B**
 - **Tranziție:B->A**



Fig.2.1. Andrei Markov, matematician rus cunoscut pentru studiul proceselor stocastice

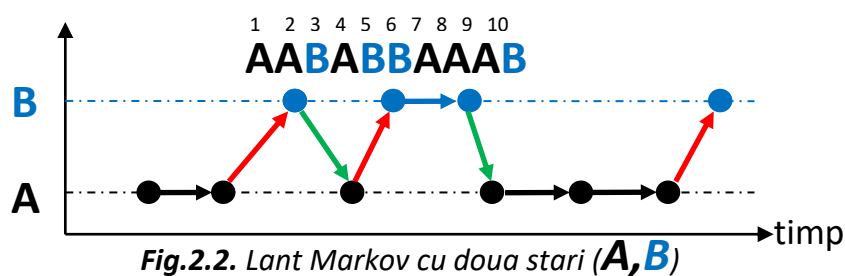


Fig.2.2. Lanț Markov cu două stări (**A,B**)

- Numărul total de tranziții ($N_{tr}=3+3+1+2=9$) (**Tabelul 2.1**):
 - $n_{AA}=3$
 - $n_{AB}=3$
 - $n_{BB}=1$
 - $n_{BA}=2$
- Matricea probabilităților de tranziție (MT) (**Tabelul 2.2**)
- Diagrama matricii probabilităților de tranziție (**Fig.2.3**)

Tabel 2.1. Număr de tranziții

		A	B	Număr total pe rând (NTR)
NT:	A	3	3	6
	B	1	2	3

Tabel 2.2. Matricea probabilităților de tranziție

		A	B	NTR
MT:	A	3/6	3/6	1
	B	1/3	2/3	1

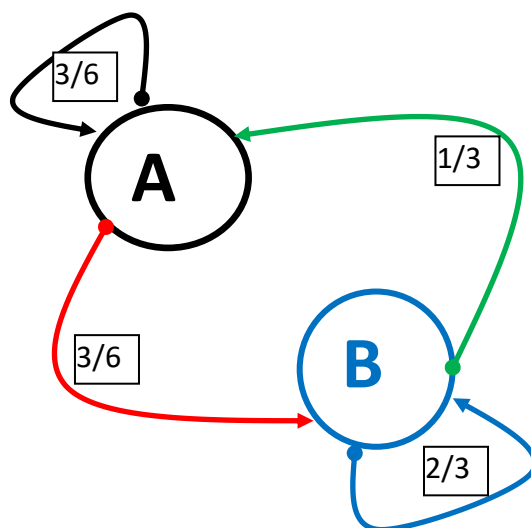


Fig.2.3. Diagrama matricii probabilităților de tranziție ($MT(A,B)$)

Incertitudinea asociată matricii probabilităților de tranziție este măsurată prin **entropia matricii probabilităților de tranziție** ($H(MT)$) calculată cu relația:

$$H(MT) = -\left(p_A \cdot (p_{AA} \cdot \log_2(p_{AA}) + p_{AB} \cdot \log_2(p_{AB})) + p_B \cdot (p_{BA} \cdot \log_2(p_{BA}) + p_{BB} \cdot \log_2(p_{BB}))\right)$$

care aplicată pentru cele 4 tranziții (**Fig.2.2** și **Tabelul 2.2**) este:

$$H(MT) = -\left(\frac{6}{10} \cdot \left(\frac{3}{6} \cdot \log_2\left(\frac{3}{6}\right) + \frac{3}{6} \cdot \log_2\left(\frac{3}{6}\right)\right) + \frac{4}{10} \cdot \left(\frac{1}{3} \cdot \log_2\left(\frac{1}{3}\right) + \frac{2}{3} \cdot \log_2\left(\frac{2}{3}\right)\right)\right) = 0,97$$

sau în forma generală pentru N_s stări este:

$$H(MT) = -\sum_{i=1}^{i=N_s} p_i \cdot \sum_{j=1}^{j=N_s} p_{ij} \cdot \log_2(p_{ij})$$

cu valorile extreme:

- $H(MT) = 0$: atunci când procesul se află într-o singură stare și **cunosterea evoluției procesului este completă**.
- $H(MT) = \log_2 2 = 1$: atunci când toate probabilitățile din matricea de tranziție sunt egale între ele și **incertitudinea asupra stării în care va trece procesul este maximă**.

1.1.2.2. Aplicație 1D

Să se calculeze **entropia matricii de tranziție** pentru un punct de observație dintr-o rețea de monitorizare, pe baza unei serii de 50 de valori măsurate la **interval de timp egal** ($\Delta t = 7 \text{ zile}$)(**Tabelul 2.3**).

Tabelul 2.3. Serie de valori măsurate într-un punct de observație al rețelei de monitorizare

Nr	V(T)	Nr	V(T)	Nr	V(T)	Nr	V(T)
1	17.07	14	17.5	27	25.29	40	20.48
2	19.06	15	28.04	28	28.12	41	25.02
3	17.75	16	20.48	29	31.94	42	26.56
4	25.34	17	28.55	30	19.89	43	26.91
5	19.96	18	17.32	31	28.38	44	17.06
6	30.23	19	19.18	32	18.11	45	26.99
7	20.88	20	30.49	33	23.51	46	32.91
8	27.37	21	18.05	34	17.16	47	28.85
9	31.37	22	27.25	35	25.14	48	24.98
10	17.3	23	32.64	36	28.64	49	26.52
11	24.23	24	26.16	37	32.84	50	30.28
12	26.46	25	26.02	38	31.25		
13	29.91	26	32.24	39	18.29		

Rezolvare 1D:(ENTROPIE_1D.xls)

Etapile prelucrărilor pentru calculul **entropiei matricii de tranziție** sunt:

- definirea grupelor valorice, divizând amplitudinea de variație în intervale egale (**Fig.2.4.**); pentru cele $N=50$ de valori disponibile (**Tabelul 2.3**) se împarte amplitudinea selecției în trei grupe valorice (A,B,C; $\Delta = 5,28$):
 - $A \in [V \text{ min} = 17,07; V \text{ min} + \Delta = 22,35)$
 - $B \in [22,35; 27,63)$
 - $C \in [27,63; 32,91]$

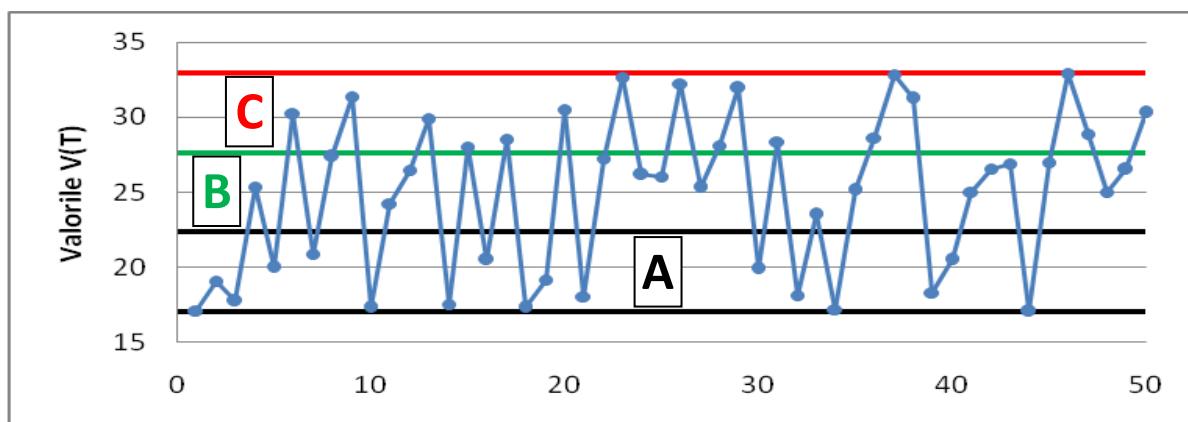


Fig.2.4. Seria de 50 de valori ale variabilei V (măsurate la un interval: $\Delta t = 7 \text{ zile}$)

NOTA. Numărul grupelor valorice se **alege** în funcție de gradul de detaliu la care se dorește cunoașterea variației în timp a variabilei în perioada monitorizată.

- Reprezentarea grafică, pentru verificarea grupării datelor (**Fig.2.4**):
 - variația în timp a variabilei;
 - limitele de separație pentru grupele valorice definite.
- Calculul frecvențelor absolute pentru cele trei grupe valorice:
 - $n_A = 17$
 - $n_B = 16$
 - $n_C = 17$
- Calculul frecvențelor relative (probabilităților medii) pentru cele trei grupe valorice:
 - **A:** $p_A = \frac{n_A}{N} = \frac{17}{50} = 0,34$
 - **B:** $p_B = \frac{n_B}{N} = \frac{16}{50} = 0,32$
 - **C:** $p_C = \frac{n_C}{N} = \frac{17}{50} = 0,34$
- Stabilirea succesiunii celor 50 de valori codificate după încadrarea în cele trei grupe valorice (**Tabelul 2.4**)

Tabelul 2.4. Succesiunea celor 50 de valori, codificate A,B,C.

Nr	V(T)	Nr	V(T)	Nr	V(T)	Nr	V(T)
1	A	14	A	27	B	40	A
2	A	15	C	28	C	41	B
3	A	16	A	29	C	42	B
4	B	17	C	30	A	43	B
5	A	18	A	31	C	44	A
6	C	19	A	32	A	45	B
7	A	20	C	33	B	46	C
8	B	21	A	34	A	47	C
9	C	22	B	35	B	48	B
10	A	23	C	36	C	49	B
11	B	24	B	37	C	50	C
12	B	25	B	38	C		
13	C	26	C	39	A		

- Calculul frecvenței absolute a celor nouă tipuri de tranziții posibile (**Tabelul.2.5**)

Tabel 2.5. Număr de tranziții

		A	B	C	Număr total pe rând
NT:	A	4	8	5	17
	B	3	5	8	16
	C	9	3	4	16

- Calculul matricii probabilităților de tranziție (**Tabelul.2.6**)

Tabel 2.6. Matricea probabilităților de tranziție

		A	B	C	Suma probabilităților pe rând
NT:	A	0.24	0.47	0.29	1
	B	0.19	0.31	0.50	1
	C	0.56	0.19	0.25	1

- Realizarea diagramei matricii probabilităților de tranziție (**Fig.2.5**)

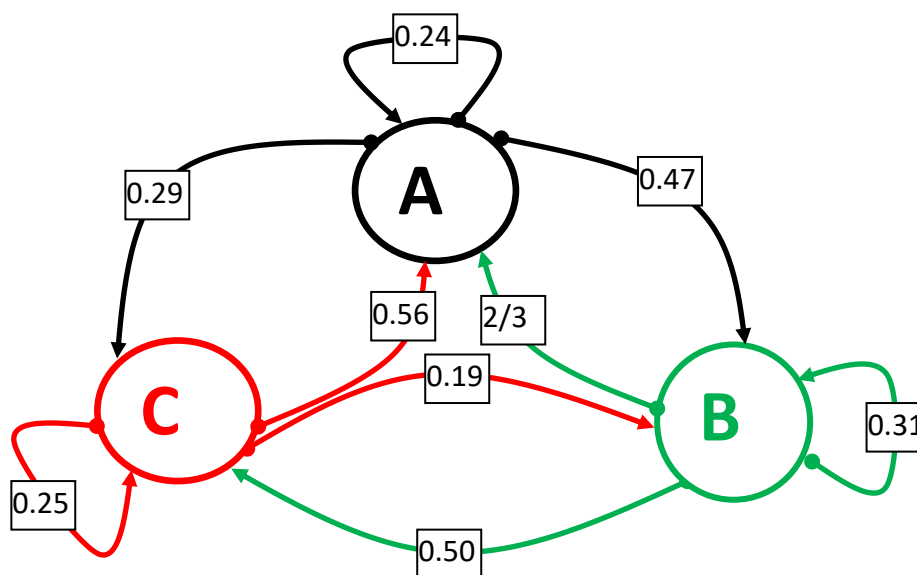


Fig.2.5. Diagrama matricii probabilităților de tranziție pentru trei stări:

A, B, C

- Calculul **entropiei matricii probabilităților de tranziție** (**Tabelul 2.6**) a succesiunii celor 50 de valori, codificate A,B,C ($N_s = 3$; **Tabelul 2.4**):

$$H(MT) = - \sum_{i=1}^{i=N_s} p_i \cdot \sum_{j=1}^{j=N_s} p_{ij} \cdot \log_2(p_{ij}) = 1,47$$

1.1.2.3. Concluzii 1D

În raport cu entropia maximă a matricii probabilităților de tranziție pentru un lanț Markov cu 3 stări distincte ($H_{MAX}(A, B, C) = \log_2 3 = 1,58$), entropia $H(MT) = 1,47$ indică un **grad de incertitudine ridicat** privind starea în care va trece procesul dintr-o stare cunoscută. O astfel de situație recomandă pentru variabila studiată **reducerea intervalului de timp (Δt)** dintre două măsurători consecutive.

Se recomandă calculul matricii de tranziție pentru toate punctele de observație ale rețelei de monitorizare și identificarea zonelor în care este necesară modificarea intervalului de timp dintre două msurători consecutive, în scopul creșterii gradului de cunoaștere a evoluției în timp a variabilei monitorizate.

1.2. Metoda punctului fictiv

Metoda punctului fictiv trebuie să stabilească punctele de observație **suplimentare** care asigură reducerea abaterii standard de estimare prin kriging (Kriging Standar Deviation), pe toată suprafața investigată, sub valoarea maximă admisă, precizată drept criteriu de valabilitate a rezultatelor investigației.

[Metoda punctului fictiv : “Scrădeanu, D, Popa, R.-GEOSTATISTICA APLICATA, 2001, EUB)

1.2.1. Noțiuni elementare

Reducerea erorii de estimare zonală se poate realiza doar prin îndesirea 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_o)) asociat unui punct de estimare (p_o) prin introducerea unui *punct fictiv* în zona lui de influență se estimează cu relația:

$$CP(p_o) = \frac{(\tilde{\sigma}_R^2)_f - (\tilde{\sigma}_R^2)_i}{(\tilde{\sigma}_R^2)_f}$$

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.

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* 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 *hărților 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.

1.2.2. Etapele de aplicare ale metodei punctului fictiv

Etapele principale ale metodei punctului fictiv sunt:

- Analiza variografică a variabilei la momentul inițial: $V(T_o)$

- Calculul variogramei de suprafață
 - Estimarea caracteristicilor structurii spațiale:
 - Modelul de variogramă
 - Parametrii de anizotropie
 - Raportul de anizotropie (R/r)
 - Orientarea direcției de continuitate maximă(θ)
 - Estimarea distribuției abaterii standard de interpolare (KSD)
 - Identificarea zonelor cu abateri standard de interpolare mai mari decât KSD maxim admis.
 - Stabilirea poziției punctelor “fictive” necesare reducerii KSD sub valoarea maximă admisă
 - Recalcularea KSD după introducerea punctelor “fictive”.
 - Definitivarea rețelei de monitorizare.
 - Calculul și reprezentarea grafică a distribuției erorilor de estimare pentru rețeaua de investigare definitivată, pentru un $\alpha = 10\%$, N și KSD utilizând funcția din excel:
 - $\varepsilon(x_i, y_j, \alpha) = CONFIDENCE(\alpha, KSD(x_i, y_j), N)$
- în care (x_i, y_j) sunt coordonatele nodurilor rețelei de interpolare pentru KSD

Finalizarea aplicării **metodei punctului fictiv** se ilustrează cu:

- harta cu poziția punctelor de investigare inițiale;
- harta distribuției KSD inițiale cu evidențierea zonelor în care este depășită KSD maximă acceptată;
- harta cu poziția **punctelor fictive** care reduc KSD sub valoarea maximă acceptată
- harta cu distribuția **valorilor variabilei** $V(x_i, y_i)$ și a **erorilor de estimare** ($\varepsilon(x_i, y_j, \alpha)$) pentru rețeaua finală de investigare, completată cu **punctele fictive**.

1.2.3. Aplicație 2D

Să se optimizeze rețeaua de explorare hidrogeologică a acviferului sub presiune din culcușul stratului V de lignit, în interfluviul Motru-Jiu.

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 hidrodinamice.

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.

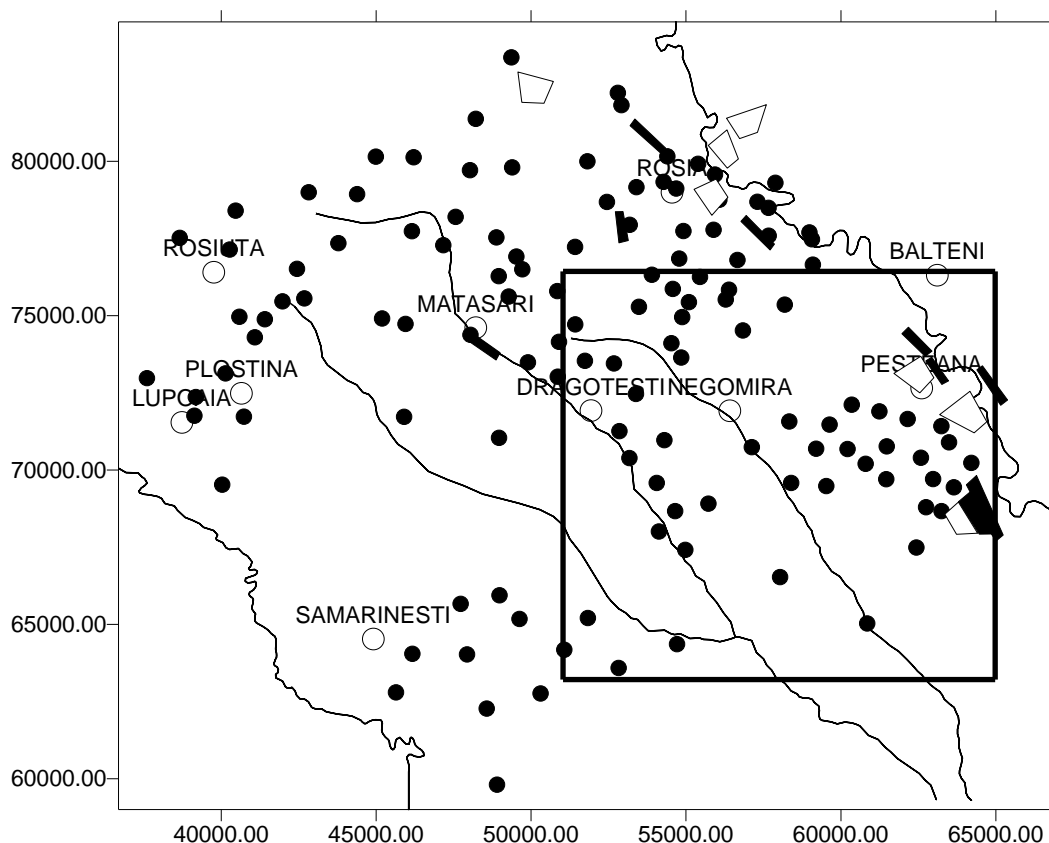


Fig.3.14. Amplasarea forajelor hidrogeologice de

cercetare a orizonturilor acvifere
 - sectorul selectat pentru aplicatie.

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 nerepresentative 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 omnidireționale* 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: $r = 7000 \text{ m}$.

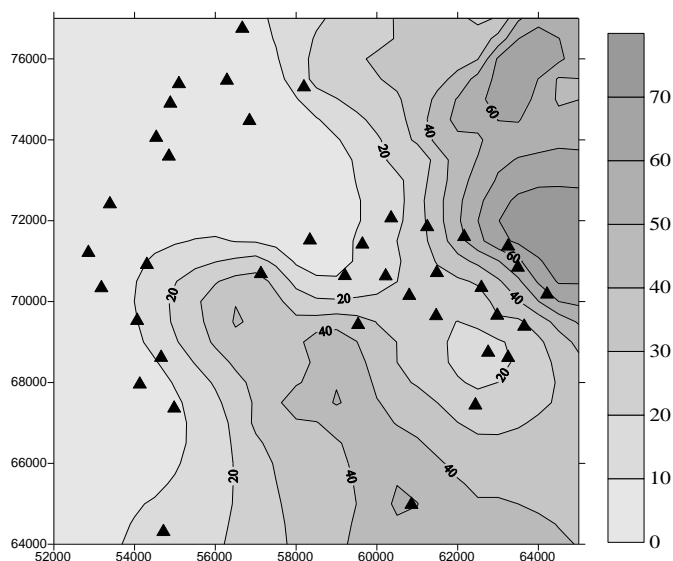


Fig.3.15. Distribuția transmisivității acviferului din culcușul stratului V calculată cu valori normalizate.

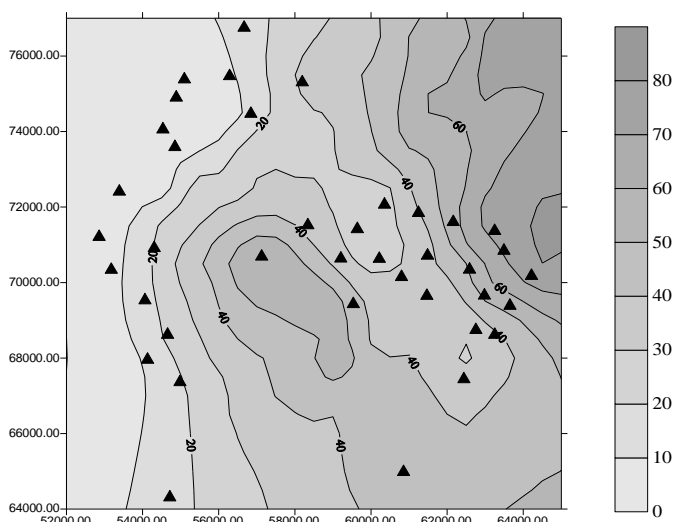


Fig.3.16. Distribuția transmisivității acviferului din culcușul stratului V calculată cu valori nenormalizate.

discretizare.

Harta conturală a *abaterei 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 $2 \text{ m}^2/\text{zi}$.

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, nelogaritmăte (Fig.3.16).

Diferența dintre cele două estimări (Fig.3.17), cea *incorectă*, realizată cu valorile originale (cu distribuție logaritmă; 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 logaritmăte (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

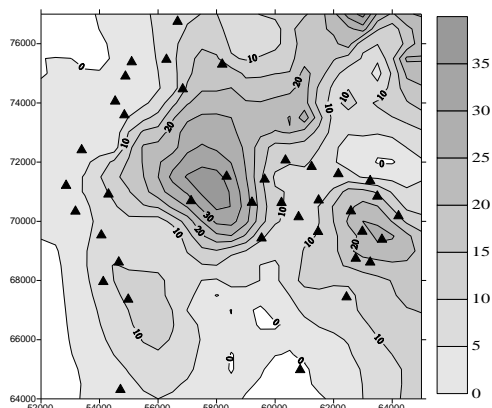


Fig.3.17. Distribuția supraestimărilor transmisivității acviferului din culcusul stratului V

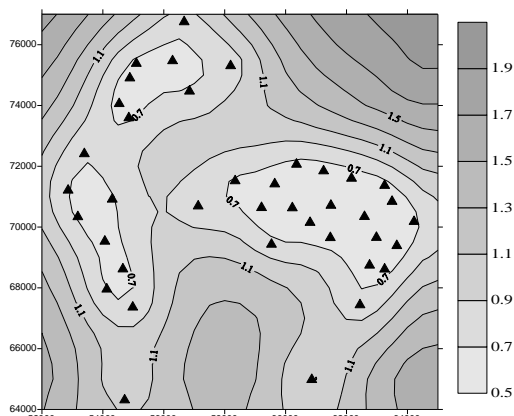


Fig.3.18. Distribuția abaterii standard de estimare a transmisivității acviferului din culcusul stratului V

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.

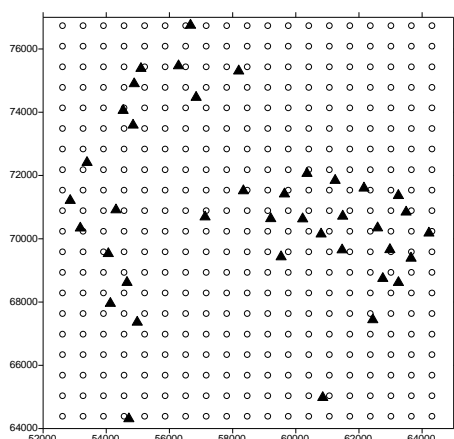


Fig.3.19. Distribuția punctelor fictive (o) și a forajelor hidroaeroaice (▲) din perimetru.

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**).

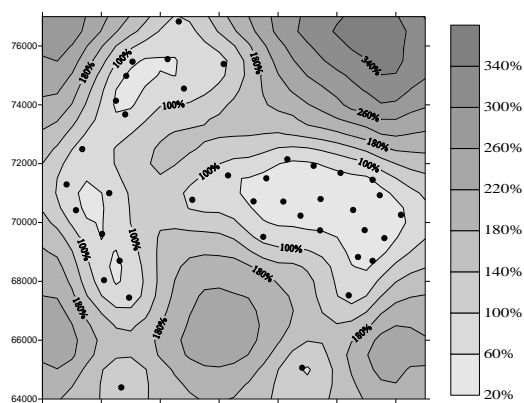


Fig.3.20 Distribuția câștigului de precizie determinat de introducerea celor 380 de puncte fictive

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 *abateri 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"!*

BIBLIOGRAFIE

1. Keith Conrad, Probability distributions and maximum entropy (<http://www.math.uconn.edu/~kconrad/blurbs/analysis/entropypost.pdf>)
2. Scădeanu Daniel, Popa Roxana, [2001, 2003], Geostatistică aplicată, Editura Universității din București
3. Scădeanu Daniel, [1995], Informatică geologică, Editura Universității din București