

Geofizikai adatsorok szegmentálása

PUSZTA SÁNDOR*

A geofizikai adatsorok feldolgozásában, az értelmezés előkészítésében fontos lépés a lényeges változások helyeinek kijelölése és ezen helyek között az adatok állandó értékkel helyettesítése, az eredeti adatsor lépcsős függvény közelítése. A következő módszereket vizsgáltuk meg: Walsh-transzformáció, átlagképzés, mediánszűrés és a bemenetet Markov-folyamatnak tekintő szegmentálás.

Az eljárások két nagy csoportba oszthatók:

- a lehetséges intervallumok határai bizonyos értelemben rögzítettek, az intervallumon belüli érték határozandó meg (példa a Walsh-tartományban végzett aluláteresztő szűrés),
- az intervallumokban felvett értékek száma költől, az intervallumok határai az eljárásból adódnak (példa: Markov folyamat alapján végzett szegmentálás)

Az eljárásokat akusztikus karotázisadatokból számított sebesség adatokon hasonlítjuk össze.

Egyik eljárás sem tekinthető tökéletesnek. Valószínűleg mindig szükséges lesz a kísérleti feldolgozás és a kiértékelő döntése, esetleg több eljárás egyesítése.

A módszerek előnye, hogy szubjektív hibáktól mentes és gyors szegmentálást biztosítanak.

В обработке геофизических данных и проведения их интерпретации важным шагом является выделение участков резкого изменения значений данных; аппроксимация постоянными значениями данных находящихся между выделенными участками, т. е. подготовка ступенчатой модели ряда геофизических данных. В работе проведено исследование следующих методов: трансформация Walsh, осреднение, медианная фильтрация и сегментация входных данных рассматриваемых как процесс Маркова.

Все методы можно разделить на две большие группы:

- возможные границы интервалов в определенном смысле фиксированы, в пределах интервала определены его характеристики (пример- в области Walsh проведенная фильтрация нижних частот),
- в интервалах число взятых значений ограничено, границы интервалов зависят от применяемого метода (напр имерпроведенная сегментация на основе принципа Маркова).

Методы сравниваются между собой со значениями скоростей полученных по данным акустического каротажа.

У каждого из методов есть свои недостатки. Скорее всего всегда нужно будет прибегать к пробной обработке или применять сразу несколько методов.

Преимущество методов – гарантируют избежать от субъективных ошибок вместе с тем быструю сегментацию.

The indication of the places of significant changes and the substitution of the measured data between these points by constant values producing the box-car function representation of the original data series are important steps both in the processing of the geophysical data series and in the preparation of the interpretation. The following methods have been examined: the Walsh-transformation, the method of averages, the median-filtering and the segmentation method considering the input data as a Markov process.

The following two groups of the methods can be determined:

- the boundaries of the possible intervals are fixed in a sense and the constant value inside each interval is to be determined. (An example is the low-pass filtering in Walsh domain.)
- the number of the values is determined in the intervals and the boundaries are produced by the method. (An example can be the Markov process type of segmentation.)

The techniques are compared using velocity data calculated from acoustic logs. None of the methods can be considered perfect. Likely the experimental processing phase the decision of the interpreter and the combined application of some of the methods will be needed.

The advantage of these methods is, that a quick segmentation can be completed which is free from subjective errors.

* ELTE Geofizikai Tanszék, Budapest

Bevezetés

Geofizikai adatok feldolgozásában lényeges feladat lehet a mért mennyiség osztályokba sorolása. A szeizmikus értelmezés automatizálásának egyik eleme a szeizmikus csatorna vagy a szeizmikus időszelvény bizonyos szempontból eltérő részeinek elkülönítése. Ez pedig különböző tulajdonságok szerinti osztályokba sorolásra és a többdimenziós osztályozás alapján meghozott döntésekre épülhet. Az osztályokba sorolás helyett gyakran a szegmentálás kifejezést használják. Szemléletesen fogalmazva az eljárást *lépcsős függvénnyel való közelítésnek* is nevezhetjük.

A szegmentálás több módszerrel is elvégezhető. Ezek közül eléggé kézenfekvő a Walsh-transzformáció alkalmazása, mert definíciója miatt szakaszonként állandó függvényeket állít elő, melyek bizonyos értelemben közelítik az eredeti függvényt. De elvégezhető a szegmentálás átlagképzésekkel, mediánszűrővel és a szegmentálandó függvényt Markov-folyamatnak tekintő eljárással is.

A dolgozat célja ezen eljárások ismertetése és összehasonlítása. Az illusztráció egy Battonya környéki mélyfúrásban meghatározott sebességszelvény (CVL) mintavételezett adatarendszer. Markov-folyamatot felhasználó módszert írt le *Vérmes (1984)*. A Walsh-transzformáció alkalmazásával a magyar szakirodalom még nem foglalkozott. Emiatt ezt a módszert valamivel részletesebben tárgyaljuk.

Walsh-függvények

A Walsh-függvények definiálása előtt célszerű felidézni az ortogonális függvényrendszerek néhány általános tulajdonságát.

Négyzetesen integrálható függvények terében (az ún. L_2 térben) léteznek teljes ortogonális függvényrendszerek (*Szőkefalvi-Nagy, 1954*). A megszámlálható sok elemből álló rendszert szimbólikusan $\{\varphi_n\}$ -nel jelöljük. Az *ortogonalitás*, ismert módon:

$$\int \varphi_n(x) \cdot \varphi_m(x) dx = k \quad n = m \\ = 0 \quad n \neq m \quad (1)$$

Amikor $k = 1$ tetszőleges n értékre, a $\{\varphi_n\}$ függvényrendszer *ortonormált*.

Teljes ortonormált függvényrendszer szerint tetszőleges $f(x) \in L_2$ sorbafejthető:

$$f(x) = \sum_{n=1}^{\infty} C_n \cdot \varphi_n(x), \quad (2)$$

melyben a C_n együtthatók az $f(x)$ függvény $\{\varphi_n\}$ rendszer szerinti együtthatói

$$C_n = \int f(x) \cdot \varphi_n(x) dx. \quad (3)$$

Amikor a $\{\varphi_n\}$ rendszer teljes ortogonális rendszer, de nem normált az együtthatók értéke:

$$C_n = \frac{\int f(x) \varphi_n(x) dx}{\int \varphi_n^2(x) dx}. \quad (4)$$

A matematikai fizikában igen sok különböző ortogonális függvényrendszert használunk. Az alapfüggvényeket a feladat természetéhez illeszkedően választják. Ismert példák a komplex Fourier sorok és a Legendre polinomok.

A komplex Fourier sor a $\varphi_n = e^{jnx}$ rendszer szerinti sorfejtés; azaz

$$f(x) = \sum_n C_n \cdot e^{jnx} \quad (5)$$

melyben az együtthatók:

$$C_n = \frac{1}{2\pi} \int f(x) e^{-jnx} dx \quad (6)$$

A Legendre polinomok az $1, x, x^2, \dots, x^n, \dots$ függvényekből Schmidt-féle ortogonalizálási eljárással (ld. *Szőkefalvi-Nagy, 1954*) levezetett, a $[-1, 1]$ intervallumon normált rendszert alkotó polinomok. A rendszer általános tagja:

$$\varphi_m(x) = P_m(x) = \frac{1}{m!2^m} \cdot \frac{d^m}{dx^m} \cdot (x^2 - 1)^m. \quad (7)$$

Az első néhány Legendre polinom:

$$\begin{aligned} P_0(x) &= 1, & P_1(x) &= x, \\ P_2(x) &= \frac{1}{2}(3x^2 - 1), & P_3(x) &= \frac{1}{2}(5x^3 - 3x), \\ P_4(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3), & P_5(x) &= \frac{1}{8}(63x^5 - 70x^3 + 15x). \end{aligned}$$

A Legendre polinomokkal a $[-1, 1]$ intervallumban az $f(x)$ függvény

$$f(x) = \sum_{n=0}^{\infty} C_n \cdot P_n(x) \quad (8)$$

alakban állítható elő, a C_n együtthatók a

$$C_n = \frac{2n+1}{2} \int_{-1}^1 f(x) \cdot P_n(x) dx \quad (9)$$

képlettel határozhatók meg.

További példák lehetnek adott súlyfüggvényre vonatkozólag különböző intervallumokban ortogonális polinomok: Csebisev-polinomok, Jacobi-polinomok, Hermit-polinomok, Laguerre-polinomok. Jelentőségüket az adja, hogy mindegyikük egy-egy másodrendű lineáris differenciálegyenletnek tesz eleget. A differenciálegyenletek gyakran szerepelnek fizikai jelenségek és folyamatok leírásában.

A dolgozat célkitűzése szempontjából fontosabbak számunkra a lépcsős alapfüggvényekből álló ortogonális rendszerek. Az első ilyen rendszert Haar

Alfréd közölte 1909-ben doktori disszertációjában. A függvények $[0,1]$ szakaszon vannak értelmezve. Néhány kis indexű Haar-függvényt az 1. ábra mutat be. Az alapfüggvények két indexesek; az első néhány alapfüggvény definíciója:

$$X_0^{(0)}(x) = 1 \quad 0 \leq x \leq 1 \quad (10a)$$

$$X_0^{(1)}(x) = \begin{cases} 1 & \text{ha } 0 \leq x < \frac{1}{2} \\ 0 & \text{ha } x = \frac{1}{2} \\ -1 & \text{ha } \frac{1}{2} < x \leq 1 \end{cases} \quad (10b)$$

$$X_1^{(1)}(x) = \begin{cases} \sqrt{2} & \text{ha } \frac{1}{2} \leq x < \frac{3}{4} \\ -\sqrt{2} & \text{ha } \frac{3}{4} < x \leq 1 \\ 0 & \text{minden más pontban,} \end{cases} \quad (10c)$$

$$X_1^{(2)}(x) = \begin{cases} \sqrt{2} & \text{ha } \frac{1}{2} \leq x < \frac{3}{4} \\ -\sqrt{2} & \text{ha } \frac{3}{4} < x \leq 1 \\ 0 & \text{minden más pontban,} \end{cases} \quad (10d)$$

és általában (alsó index: $n \geq 0$, míg a felső index $k: 1 \leq k \leq 2^n$):

$$X_n^{(k)}(x) = \begin{cases} \sqrt{2^n} & \text{ha } \frac{k-1}{2^n} \leq x < \frac{k}{2^n} \\ -\sqrt{2^n} & \text{ha } \frac{k-1}{2^n} < x \leq \frac{k}{2^n} \\ 0 & \text{minden más pontban.} \end{cases} \quad (11)$$

Az alapfüggvények szerkezete miatt a rendszer ortogonális. Amikor az alsó indexek megegyeznek, már a $X_n^{(l)} \cdot X_n^{(k)}$ szorzat is zérus – ld. pl. az 1. ábrán $X_1^{(1)}(x)$ és $X_1^{(2)}(x)$. Amikor az alsó indexek különböznek a nagyobb alsó indexű Haar-függvény benne foglaltatik egy olyan intervallumban, melyben a kisebb alsó indexű állandó, pl. $m > n$ és akár $k \neq 1$, akár $k = 1$ esetén:

$$\int_0^1 X_m^{(l)}(x) \cdot X_n^{(k)}(x) dx = \pm \sqrt{2^n} \cdot \int_0^1 X_m^{(l)}(x) dx = 0. \quad (12)$$

Könnyen belátható, hogy a rendszer normált is.

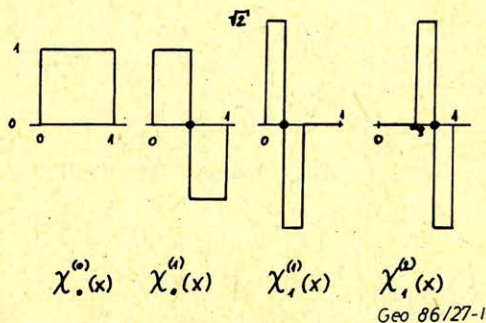
A Haar-féle rendszernél egyszerűbb a – szintén a $[0,1]$ intervallumon értelmezett – Rademacher-féle ortogonális rendszer. Ennek első néhány alapfüggvényét a 2. ábrán mutatja. A $[0,1]$ intervallumot 2^n egyenlő részre osztjuk, az $r_n(x)$ Rademacher-féle függvény az így kapott nyílt intervallumokban felváltva a $+1$ és -1 értéket veszi fel; az osztópontokban és végpontokban pedig értéke zérus. A definíciót képletekkel is megadhatjuk:

$$r_n(x) = \operatorname{sgn}(\sin 2^n \pi x), \quad (n = 1, 2, \dots), \quad (13)$$

vagy

$$r_n(x) = (-1)^{[2^n x]}, \quad (n = 1, 2, \dots). \quad (14)$$

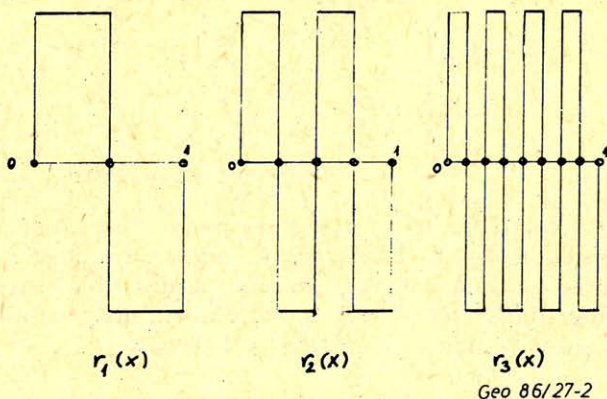
A (13) képletben sgn az előjel függvényt, a (14) képletben a szögletes zárójel az egész-rész képzést jelöli.



1. ábra. Kis indexű Haar-függvények

Рис. 1. Функция Хаара с небольшими индексами.

Fig. 1. Haar functions of low index



2. ábra. Néhány Rademacher-függvény

Рис. 2. Несколько функций Радемачера.

Fig. 2. Some Rademacher functions

A Rademacher-féle függvényrendszer ortogonális: könnyen belátható, hogy két tetszőleges függvény szorzata is $+1$ és -1 értékű szakaszokból áll és a különböző előjelű szakaszok hossza azonos, emiatt az integrál zérus. Az is nyilvánvaló, hogy a rendszer normált: tetszőleges Rademacher-függvény négyzete 1 a teljes $[0,1]$ intervallumon. A rendszer azonban *nem teljes*. Még azáltal sem tehető teljesé ha az azonosan 1 értékű függvénnyel is kiegészítjük az $r_1(x), r_2(x), \dots$, függvényrendszert. A függvények ugyanis antiszimmetrikusak a 0.5 pontra nézve:

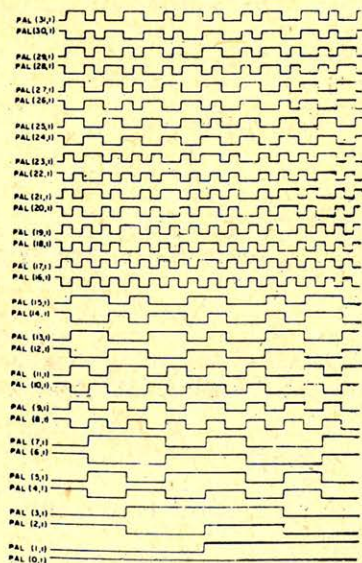
$$r_n(x) = -r_n(1-x),$$

emiatt minden olyan $f(x)$ ortogonális rájuk, amely a 0.5 középpontra nézve szimmetrikus. Ha az $f(x)$ integrálja $(0,1)$ intervallumban zérus, az $f(x)$ még a rendszerhez hozzávett, állandóan 1 értékű függvényre is ortogonális.

A Walsh-függvények rendszere a Rademacher-függvények teljes ortonormált kiterjesztésével is levezethető. A 0 indexű Rademacher-függvénynek tekintjük az állandóan 1 értékű függvényt. Az n (egész) indexű Walsh-függvényt Rademacher-függvények szorzata adja:

$$\text{WAL}(n, t) = \prod_{i=1}^m r_i(t). \quad (15)$$

Szorzatban csak azok az $r_i(t)$ függvények szerepelnek, melyekre az n index Gray kódban megadott i -edik bithelyén 1 áll. A teljesség érdekében röviden meg-



Geo 86/27-3

3. ábra. „Természetes” sorrendű Walsh-függvények

Рис. 3. Функции Уалша в естественной последовательности.

Fig. 3. Walsh functions of natural sequence

adjuk a Gray-kód definícióját. Először definiálnunk kell a *modulo* -2 összegezést: ez bináris összeadás, az átvitel elhagyásával. A lehetséges esetek:

$$0 \oplus 0 = 0, 0 \oplus 1 = 1, 1 \oplus 0 = 1 \text{ és } 1 \oplus 1 = 0. \quad (16)$$

(A közönséges összeadástól megkülönböztetendő, az összegezés jelét körben helyezzük el). Az n szám bináris alakja legyen:

$$n = (b_m b_{m-1} \dots b_1). \quad (17)$$

A Gray-kód bitjeit

1. táblázat Таблица 1. Table 1

Decimális	Binális	Gray kód
0	0000	0000
1	0001	0001
2	0010	0011
3	0011	0010
4	0100	0110
5	0101	0111
6	0110	0101
7	0111	0100
8	1000	1100
9	1001	1101
10	1010	1111
11	1011	1110
12	1100	1010
13	1101	1011
14	1110	1001
15	1111	1000

$$g_i = b_i \oplus b_{i+1} \quad (18)$$

definiálja, azaz

$$n = (g_m, g_{m-1} \dots g_1). \quad (19)$$

Az első néhány természetes szám decimális, bináris és Gray-kódú megadását az 1. táblázat tartalmazza.

A (15) előállítás illusztrálására határozzunk meg néhány kis indexű Walsh-függvényt. Ha $n = 1$, a táblázat 2. sorából láthatóan a Gray-kód első bit-helyén 1 áll, azaz

$$WAL(1, t) = r_1(t).$$

Amikor $n = 2$, a Gray-kód első és második helyén áll 1, azaz

$$WAL(2, t) = r_2(t) \cdot r_1(t).$$

A (15) alakú előállítást néha úgy módosítják, hogy a Walsh-függvények szorzata jobban hasonlítson a harmonikus függvények szorzatához. (A módosítás egyszerű fázistolás). Így jutnak a 3. ábrán bemutatott ún. természetes sorrend-be rendezett Walsh-függvények szorzatához (Beauchamp, 1975).

Számos más előállítás is lehetséges: rekurzív generálás, Hadamard mátrixokból történő előállítás, szinusz és koszinusz függvények különböző szorzatain alapuló előállítás stb. Ezek közül csak a rekurzív előállítást ismertetjük:

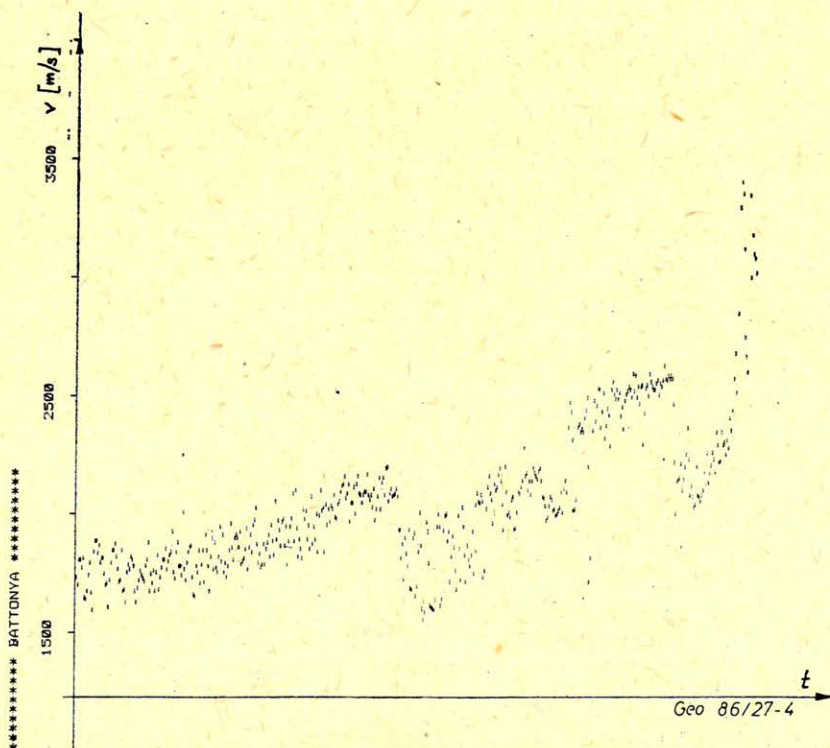
$$WAL(0, t) = 1 \quad 0 \leq t \leq 1 \quad (20a)$$

$$WAL(2j+q, t) = (-1)^{[j/2]+q} \cdot \{WAL(j, 2 \cdot t) + (-1)^{j+q} \cdot WAL(j, 2 \cdot (t-1/2))\}. \quad (20b)$$

Ha a $[0,1]$ intervallumot $N = 2^k$ egyenközű részre osztjuk, valamint a t helyett az $n = 0, 1, \dots, N-1$ indexelést használjuk:

$$WAL(2j+q, n) = (-1)^{[j/2]+q} \cdot \left\{ WAL(j, 2 \cdot n) + (-1)^{j+q} \cdot WAL(j, 2 \cdot n - N) \right\}. \quad (20c)$$

A $[0,1]$ intervallumon kívül első indexekre a WAL függvény értékét zérusnak kell tekinteni.



4. ábra. A számítási példák bemeneti függvénye

Рис. 4. Пример расчета входной функции.

Fig. 4. The input functions of the numerical examples

Átlagképzés, mediánszűrés, Markov-lánc modell

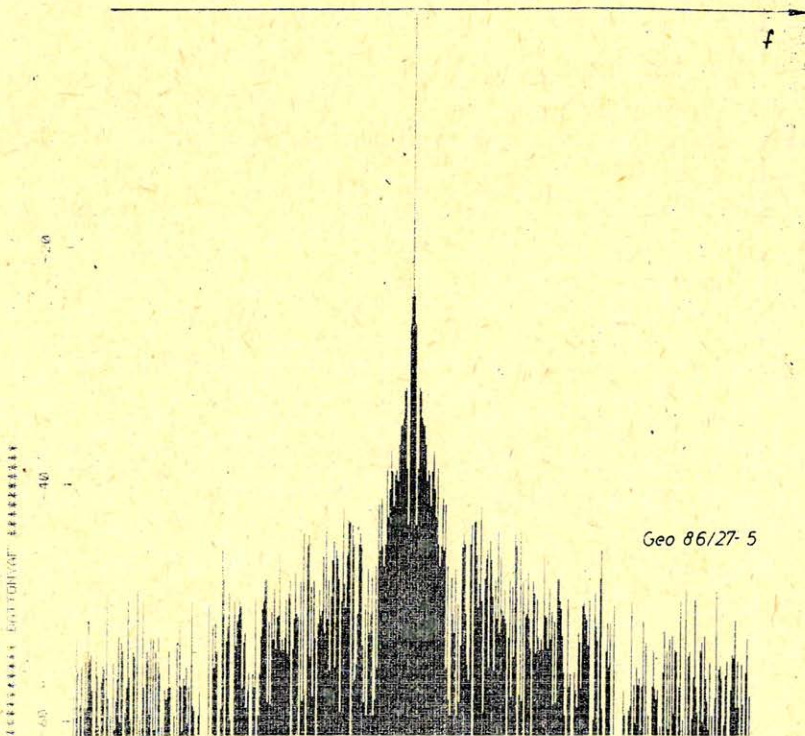
Ebben a fejezetben a nem ortogonális átalakításokról lesz szó, melyek bizonyos mértékben a szegmentálási feladat megoldásai is szolgálhatnak.

Átlagképzés rögzített intervallumokon

A bemenet adatait valamilyen megfontolás szerint intervallumokra bontjuk. A kijelölt intervallumok adataiból átlagértéket számítunk. A kimeneti pontokhoz a saját intervallumok átlagát rendeljük hozzá.

Átlagtól való eltérésen alapuló intervallumkijelölés

A kezdő elemtől mindig egy-egy újabb elemet véve, kiszámítandó az átlag értéke. Az intervallum határa ott van, ahol a következő elem az átlagtól egy adott kritérium szerint *túl*ságosan eltér. Ekkor a kimeneten az intervallum minden pontja az átlagértéket kapja és az „eltérő” elemtől indulva új intervallum feldolgozása kezdődik.

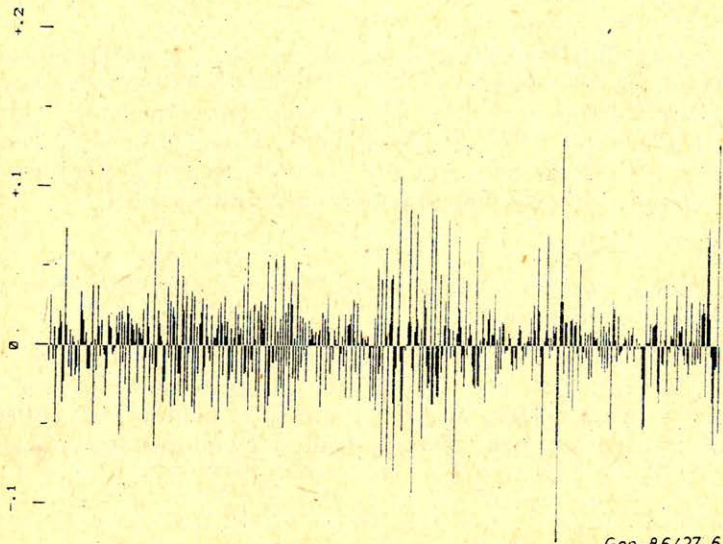


5. ábra. A bemenet Fourier transzformáltjának abszolút értéke dB skálán

Рис. 5. Абсолютное значение входной трансформации Фурье в дБ.

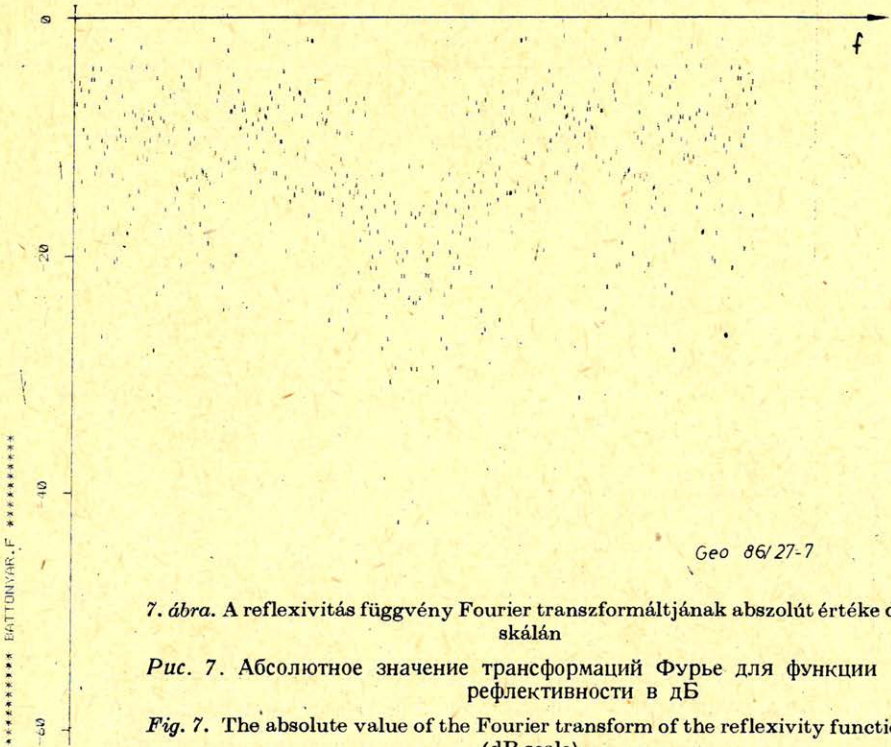
Fig. 5. Absolute value of the Fourier transform of the input (dB scale)

***** BATTONYAR ***** REFLEXIVITAS FGV *****



Geo 86/27-6

6. ábra. A bemenetből számított reflexivitas függvény
 Рус. 6. Функция рефлексивности рассчитанная по входным данным.
 Fig. 6. Reflexivity function derived from the input



Geo 86/27-7

7. ábra. A reflexivitas függvény Fourier transzformáltjának abszolút értéke dB skálán
 Рус. 7. Абсолютное значение трансформаций Фурье для функции рефлексивности в ДБ
 Fig. 7. The absolute value of the Fourier transform of the reflexivity function (dB scale)

***** BATTONYAR..F *****

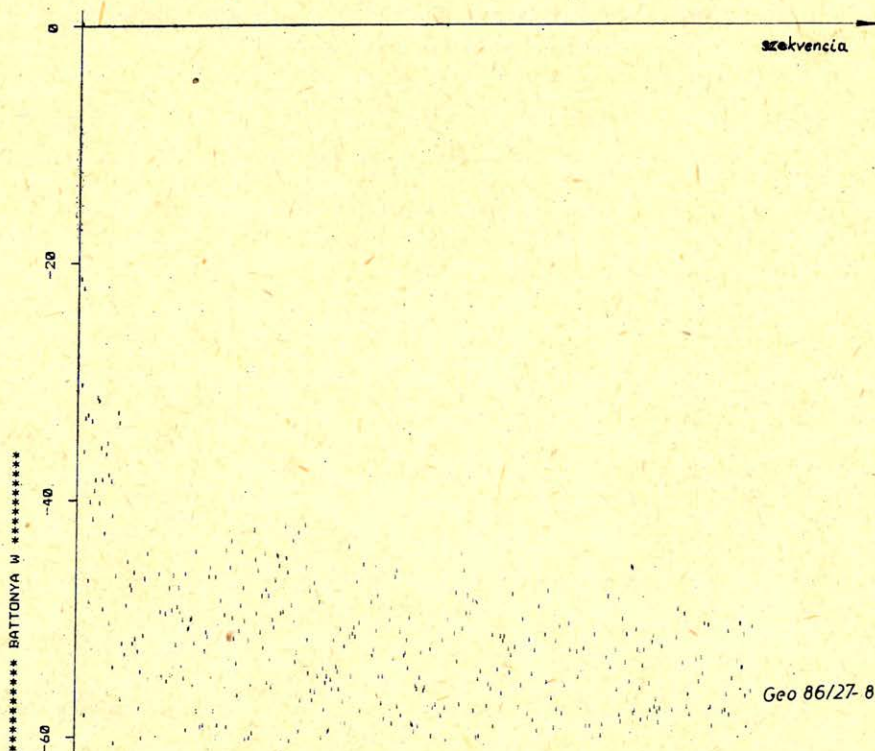
Mediánszűrés

A művelet L pontszámú futó ablakból dolgozik. Kiválasztja az aktuális, érték szerint rendezett adatok középső elemét — mediánját. Ez az érték lesz az ablak középső pontjához tartozó kimenet. Az ablak tovább lép, a rendezés és kiválasztás folytatódik. Ha az ablak a bemenet végére ért, akkor a kimenet a bemenet helyébe kerül, és a feldolgozás mindaddig folytatódik, míg változás történik az adatokon. A művelet eredménye a már stacionáriussá vált kimenet.

Markov-láncon alapuló modellezés

E témakör elméleti alapjainak leírása megtalálható a szakirodalomban, így itt csak egy rövid vázlatra szorítkozunk.

Az átalakítandó adatsort tekintsük egy véletlen folyamat realizációjának, melyben egy kapcsolatot írunk elő a valószínűségi változó aktuális értéke és az őt követő valószínűségi változó értéke között. A kapcsolat pontosabb megfogalmazása a következő: megadjuk annak valószínűségét, hogy a következő pillanat-



8. ábra. A bemenet Walsh transzformáltjának abszolút értéke dB skálán

Рис. 8. Абсолютное значение входной трансформации Уалша в дБ.

Fig. 8. The absolute value of the walsh transform of the input (dB scale)

ban x_k paramétert fogunk mérni feltéve, hogy most x_l értéket mérünk, tehát megadjuk a

$$P_r(U_{t+1} = x_k | U_t = x_l) \quad (21)$$

feltételes valószínűséget.

A problémakör kezelhetősége érdekében többféle egyszerűsítő feltevéssel élünk:

feltesszük, hogy a (21) kapcsolat nem függ t -től;

feltesszük, hogy a lehetséges paraméterek (x) értékkészlete diszkrét, végül

feltesszük, hogy a bemenő adatok halmaza is diszkrét.

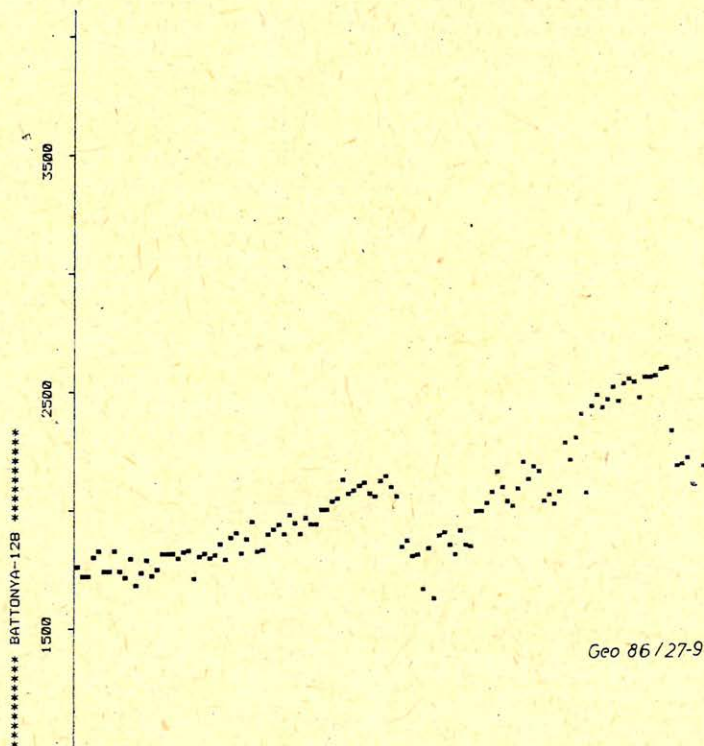
A feltételes valószínűség (21) kifejezésben emiatt $0 \leq t < N$, ahol N a mért adatsor hossza, $l \leq k, l \leq M$, ahol M a lehetséges paraméterek száma.

A valós mérési anyagon nem teljesül az $x_t \rightarrow U_t$, hanem $x_t \rightarrow U_t + n_t$, ahol n_t zaj.

A Vermes (1984) által közölt Markov-lánccal történő modellezési algoritmust alkalmaztuk a következők szerint:

A bemenetet egydimenziós Markov láncsal közelítettük;

Az átmeneti valószínűség mátrixot a jeltől függetlenül adtuk meg, minden P_{ik} ($i \neq k$) esetén azonosnak. Az $i = k$ eset valószínűsége (a paraméter ebben a lépésben nem változik) a program egyik bemenő adata.



9. ábra. Walsh tartománybeli aluláteresztő szűrés eredménye. Levágási szekvencia = 128.

Рис. 9. Результат низкочастотной фильтрации в области Уалша.
Порог среза равен 128 секвенции.

Fig. 9. The result of a low-pass filtering in Walsh domain. Cut-off sequence = 128.

A lehetséges paraméterek értékkészletét a bemeneti adatok által átfogott tartomány egyenközű beosztásával állítottuk elő. Az, hogy hány részre legyen osztva ez az intervallum, a program bemenő adata volt.

Annak érdekében, hogy a bemenethez jól illeszkedő Markov-láncot találjon a megvalósított eljárás, egy veszteségfüggvényt minimalizál.

A veszteségfüggvény komponensei:

$$D_{j,i} = \frac{1}{2\sigma^2}(u_i - x_j)^2 \quad (22)$$

u_i a mért érték az i időpontban

x_j a j -edik lehetséges állapot

σ a veszteségjárulék súlyozó tényezője (szórása).

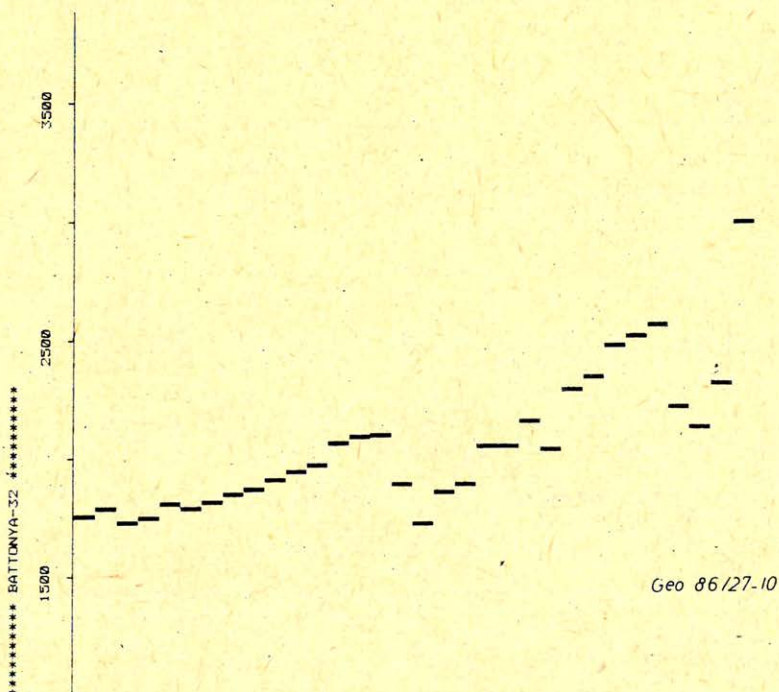
Tudjuk, hogy általában $u_i \neq x_i$. D értéke a mért jelhez közeli paraméter választásakor lesz kicsi.

$$T_{j,k} = -\ln P_{jk} \quad (23)$$

j, k a lehetséges állapotok indexei

P_{jk} a $j \rightarrow k$ átmenet valószínűsége.

T értéke nagy, ha kis valószínűségű átmenetet tételezünk fel.



10. ábra. Walsh tartománybeli aluláteresztő szűrés eredménye. Levágási szekvencia = 32.

Рис. 10. Результат низкочастотной фильтрации в области Уолша.
Порог среза равен 32 секвенции.

Fig. 10. The result of a low-pass filtering in Walsh domain. Cut-off sequence = 32.

A kimenet az egyes időpontokhoz tartozó azon x_j érték, melyre a veszteség-járulék minimális. Mindebből az következne, hogy a minimális veszteség akkor áll elő, amikor minden pillanatban a jelhez legközelebb eső és legvalószínűbb értékre váltana a kimenet. Ez azonban a szegmentálással kapcsolatos elvárásainak a gyors változások miatt nem felel meg. A problémát úgy oldhatjuk meg, hogy növeljük a $j. \rightarrow j.$ átmenet valószínűségét (λ). Ez a P_{ij} főátlójában álló elemek növelését jelenti.

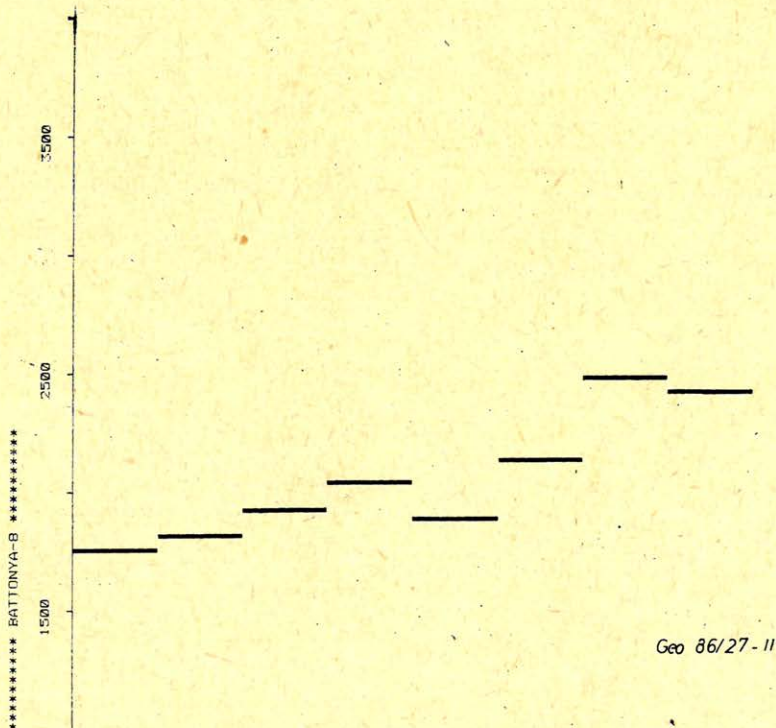
A veszteségjárulékokat minimalizáló rekurzív eljárás:

$$C_{j,0} = -\ln \alpha_i + D_{i,0} \quad (24a)$$

$$C_{j,i+1} = \min_k \{C_{k,i} + T_{k,j}\} + D_{j,i+1} \quad (24b)$$

A módszerek összehasonlítása, gyakorlati tapasztalatok

Ebben a fejezetben azokról a tapasztalatokról számolunk be, melyeket a szegmentálási feladat megoldásának különböző numerikus módszereivel kapcsolatban nyertünk.



11. ábra. Walsh tartománybeli aluláteresztő szűrés eredménye. Levágási szekvencia = 8.

Рис. 11. Результат низкочастотной фильтрации в области Уалша.
Порог среза равен 8 секвенции.

Fig. 11. The result of a low-pass filtering in Walsh domain. Cut off sequence = 8.

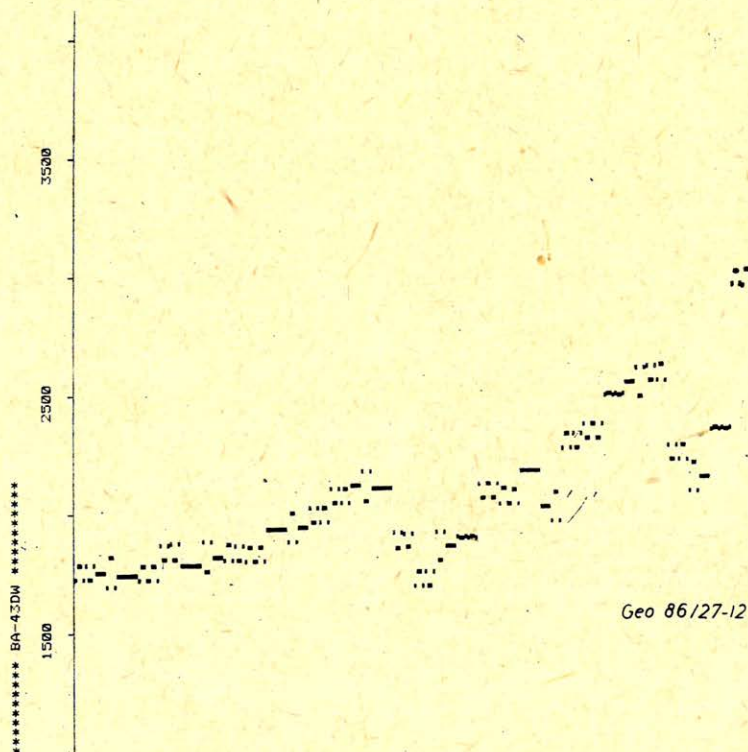
A bemutatásra kerülő számítási példánál bemenetként egy battonyai mélyfúrásban végzett akusztikus karotázsmérés adatait használtuk fel. Az akusztikus karotázsból számított intervallum-sebességeket a kétszeres vertikális terjedési időre számítottuk át és 2 ms lépésközű adatsorra ritkítettük. A kapott számsorozat 512 elemű szakaszát használtuk. Ezt mutatja be a 4. ábra. A sebességek 1500 és 3000 m/s közé esnek. Az időtengelyt a továbbiakban sem jelöljük.

Az 5. ábrán a bemeneti adatsor Fourier transzformáltjának abszolút értékét ábrázoltuk dB skálán. A 0 dB-t a maximális értékhez rendeltük, mely itt az adatsor közepén a DC komponenshez tartozik.

A 6. ábra a bemenetből számított reflexiók együtthatókat (r_k) ábrázolja

$$r_k = \frac{v_{k+1} - v_k}{v_{k+1} + v_k}$$

Az r_k értékei a -0.15 , $+0.15$ intervallumba esnek. A reflexivitás függvényből világosan látszik a vizsgált sorozat bonyolultsága. Kitűnik, hogy a szegmenztálási feladat megoldása — tehát az, hogy a bemenetet kevés, lehetőleg objektíven meghatározott intervallumra kívánjuk bontani — nem triviális, és várható-



12. ábra. Bemenettől függő Wals tartománybeli szűrés. Küszöbszint: -43 dB

Рис. 12. Фильтрация в области Уалша в зависимости от входа.
Пороговое значение: -43 дБ.

Fig. 12. Input dependent filtering in Walsh domain. Threshold level: 43 dB.

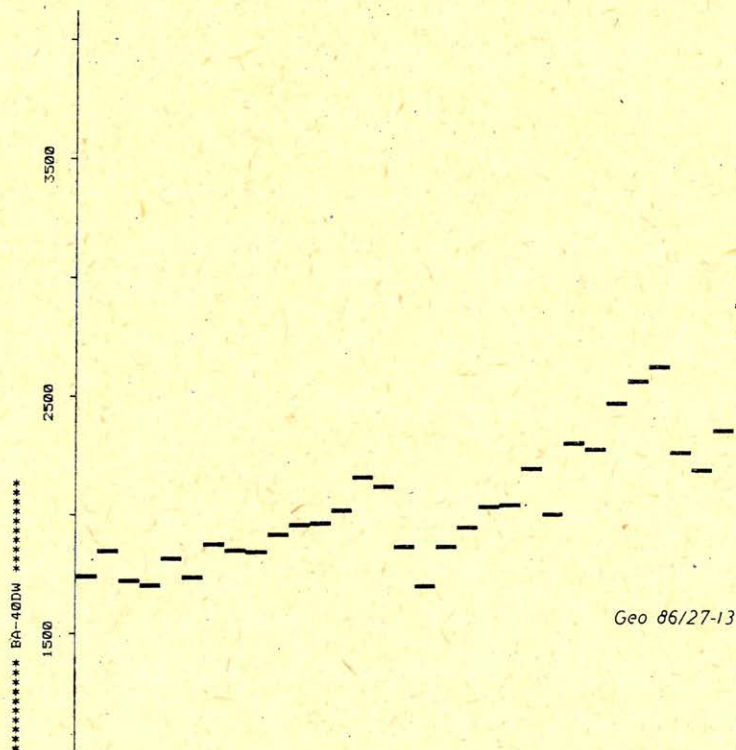
an az eredmény az alkalmazott módszerek jellegétől függően más és más lesz. Az e cikkben közölt vizsgálatok irányát éppen a különbségek feltérképezésének szándéka vezette, annak érdekében, hogy jó szeizmikus modellt konstruálhassunk.

A 7. ábrán a reflexivitás függvény Fourier transzformáltjának abszolút értéke látható, dB skálán. Ez nem fehér zaj spektruma, bár e jelleg részben a korlátozott hosszúságú bemenet miatt áll elő.

A 8. ábrán a bemenet Walsh transzformáltjának abszolút értéke látható (ugyancsak dB skálán). A nulla dB itt is a maximális értékhez van rendelve, mely ez esetben is a DC komponens. Az egyes szekvenciáértékekhez tartozó amplitúdók fentről lefelé növekvő sorrendben követik egymást.

A továbbiakban példát mutatunk Walsh-transzformáció segítségével történő átalakításra: aluláteresztő szűrést végzünk Walsh-tartományban.

A Fourier tartománynál megszokott lépéseket követjük: kiszámítjuk a függvény Walsh transzformáltját; e transzformáltat megszorozzuk egy alkalmas függvénnyel, mely esetünkben a nullától a sávhatárként kijelölt szekvenciáig 1 értékű, a nagyobb szekvenciákra 0; a szorzatot visszatranszformáljuk.



13. ábra. Bemenettől függő Walsh tartománybeli szűrés. Küszöbszint: -40 dB.

Рис. 13. Фильтрация в области Уалша в зависимости от входа.
Пороговое значение: -43 дБ.

Fig. 13. Input dependent filtering in Walsh domain. Threshold level: -40 dB.

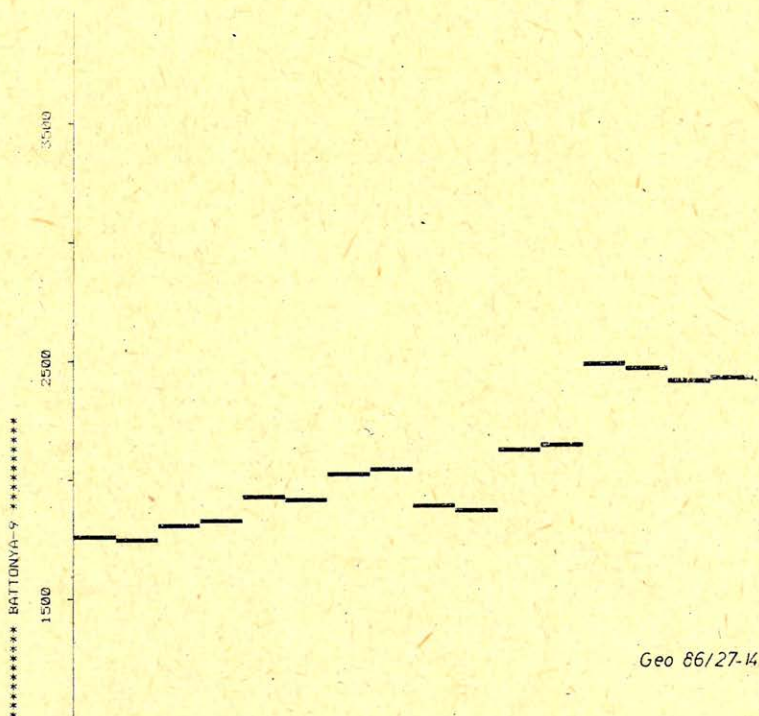
A levágási szekvencia értéke a 9., 10. és 11. ábránál sorban 128, 32 és 8 volt. Az ábrák alapján megállapíthatjuk, hogy a 128 paraméterű nem tér el lényegesen a bemenettől, a 8-as paraméterű viszont számunkra túlságosan leegyszerűsített képet ad.

Ha feltételeznénk, hogy a Walsh transzformált minden szekvenciája azonos információt hordoz, ez esetben $1/4$, $1/16$ és $1/64$ lenne a szűrt függvény információtartalma a bemenethez képest.

A bemenet Walsh transzformáltját tekintve kitűnik, hogy ez a homogenitás nem áll fenn. Ebből azonban egy ritkán használt művelethez jutunk. Az átalakítás során az információtartalmat csökkentjük úgy, hogy azt a legkevesebb Walsh transzformáltbeli komponens hordozza. Ezt úgy érhetjük el, hogy csak azokat a komponenseket tartjuk meg, melyek egy bizonyos értéknél nagyobb amplitúdójúak. A szűrő átviteli függvénye – melynek értékkészlete a 0 és 1 számok – a bemenettől függ.

A 12. és a 13. ábrák esetében használt küszöbértékek: -43 dB, illetve -40 dB.

A kedvező tulajdonságok mellett szóljunk a Walsh-függvény alkalmazásának hátrányairól is.



14. ábra. Walsh tartománybeli aluláteresztő szűrés eredménye. Levágási szekvencia = 9.

Рис. 14. Результат низкочастотной фильтрации в области Уалша.
Порог среза равен 9 секвенциям.

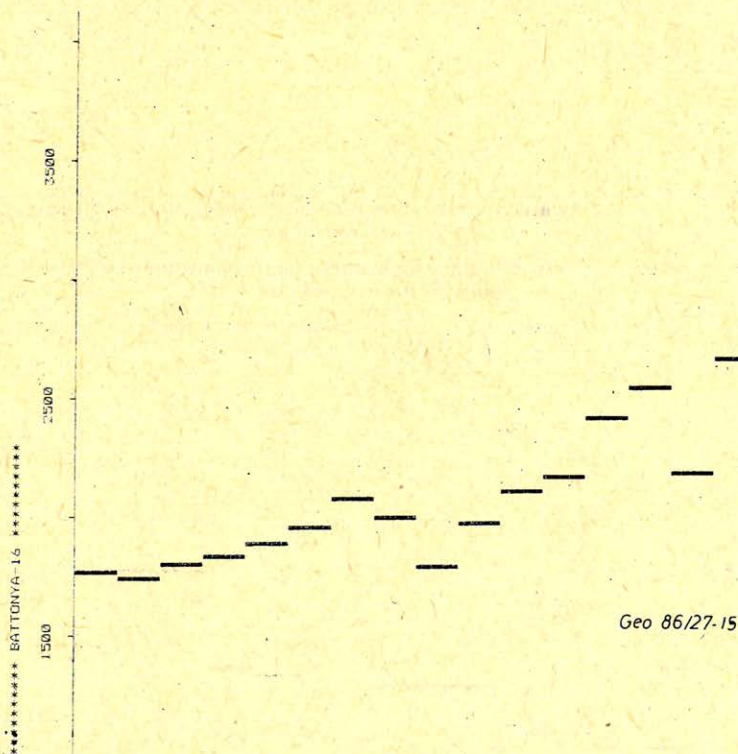
Fig. 14. The result of a low pass filtering in Walsh domain. Cut off sequence = 9.

A k indexű Walsh-függvény a teljes időtartományt k részre bontja. Az aluláteresztő szűrésnél, ha a sávhatár 2^N akkor a szűrő kimenete egy 2^N szakaszból álló függvény lesz. Ha a sávhatár $k \neq 2^N$, akkor a szűrő *felfelé kerekít*, abban az értelemben, hogy az intervallumok száma a kimeneten ismét 2 egész kitevőjű hatványa, melyre $2^N < k < 2^{N+1}$. Ha k -t növeljük, akkor az intervallumok száma nem, csak az intervallumokban felvett értékek változnak. Tehát aluláteresztő, Walsh-tartományban végzett szűréssel nem tudunk tetszőleges számú egyenközű intervallumot előállítani a kimeneten.

A 11., 14. és 15. ábra 8, 9, és 16 paraméterű Walsh-tartománybeli aluláteresztő szűrés eredményét mutatja. Itt említjük meg, hogy a 2^N paraméterű Walsh aluláteresztés művelete azonos a bemenet 2^N számú szakaszra bontásával és a szakaszokon belüli adatok átlagukkal történő helyettesítésével.

A Walsh transzformáció erősen megkötötte az intervallum hosszát. Most olyan eljárást keresünk, ahol az intervallum hossza jobban illeszkedik a bemenő jelhez.

Átlagtól való eltérésen alapuló szegmentálásban az új intervallum kijelölésének kritériuma az, hogy az új elem az előző átlagtól egy adott értéknél jobban térjen el. Ez az érték a 16. ábrán 7,5%, a 17. ábrán 10% volt. Látjuk, hogy a paraméter növelésével nagyobb szerkezeti egységeket kapunk, azonban a módszer na-



15. ábra. Walsh tartománybeli aluláteresztő szűrés eredménye. Levágási szekvencia = 16.

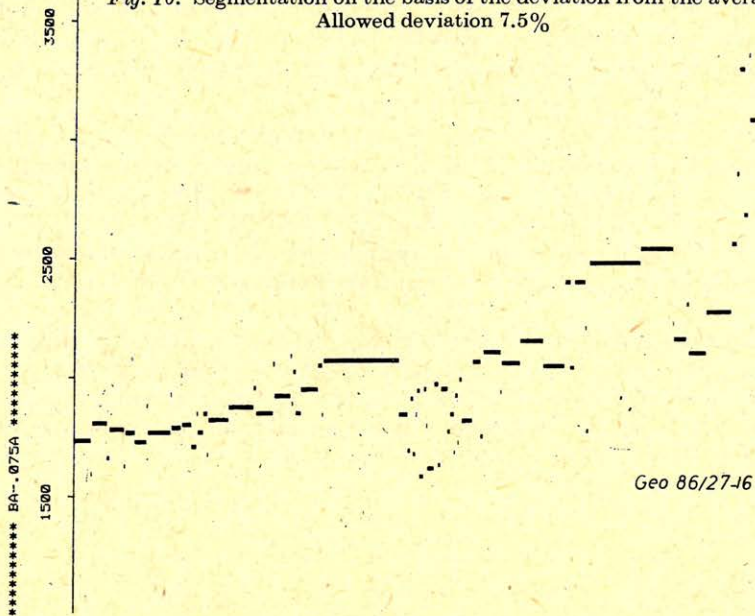
Рис. 15. Результат низкочастотной фильтрации в области Уолша.
Порог среза равен 16 секвенциям.

Fig. 15. The result of a low pass filtering in Walsh domain. Cut off sequence = 16.

16. ábra. Az átlagtól való eltérésen alapuló szegmentálás. Megengedett eltérés < 7,5%

Рис. 16. Сегментация основанная на отклонении от среднего. Допустимое отклонение < 7,5%.

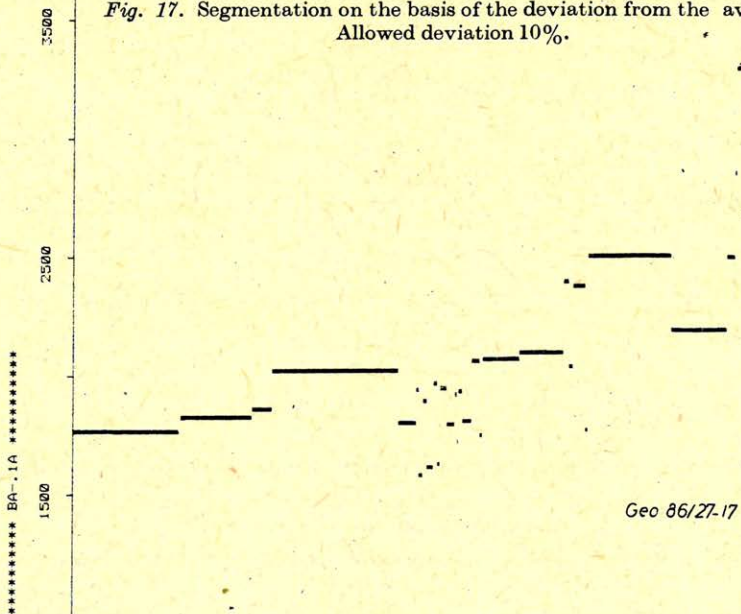
Fig. 16. Segmentation on the basis of the deviation from the average. Allowed deviation 7.5%



17. ábra. Az átlagtól való eltérésen alapuló szegmentálás. Megengedett eltérés < 10%

Рис. 17. Сегментация основанная на отклонении от среднего. Допустимое отклонение < 10%.

Fig. 17. Segmentation on the basis of the deviation from the average. Allowed deviation 10%.



gyon érzékeny a kis kiterjedésű eltérésekre is, és így az adatsor bizonyos részein a kapott kimenet ugyanolyan bonyolult, mint a bemenet.

A mediánszűrés egyetlen paramétere az ablakhossz. Ez a 18., 19. és 20. ábrán 5, 25 és 125 adat.

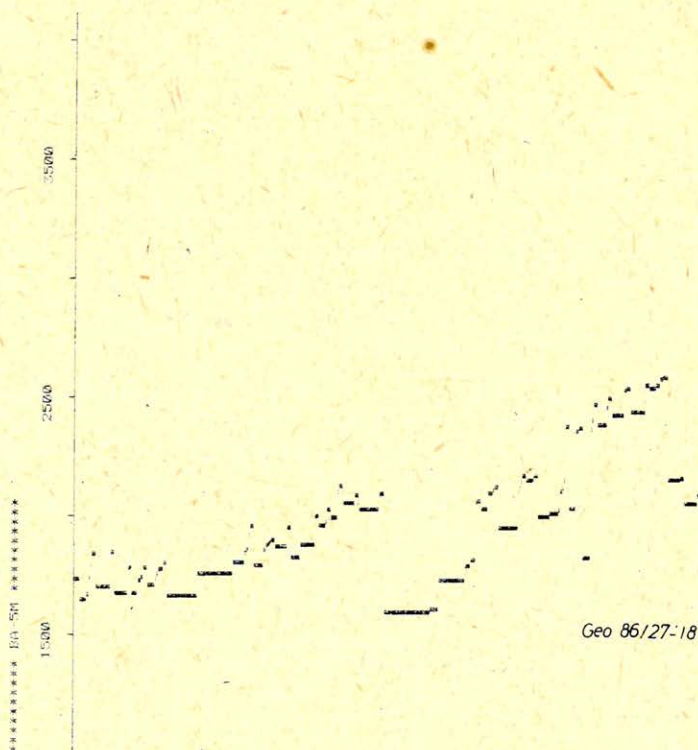
A módszerből következik, hogy az adatsor elején és végén $1/2$ ablakhossznyi rész a $1/2$ ablakhossznyi távolságban levő adat kiterjesztése lesz. A kapott eredmények érdekesek. A kapott intervallumok hosszai a bemenet szerkezetétől függően is változnak. Az ablak hosszának változtatásával pedig befolyásolni tudjuk a kimenet felbontását.

Hátrányának tekinthetjük, hogy az egyes tartományok között nem hoz létre éles átmenetet.

Az eljárás számítási időigénye az ablakhosszal négyzetesen nő.

A Markov-láncokon alapuló szegmentálást is többféle paraméterválasztással illusztráljuk.

A 21. ábrát a következő paraméterek jellemzik: 15 lehetséges szint van, a szórás értéke $\sigma = 1000$, tehát a hibajárolék D tagja igen kis súllyal szerepel, λ (főátló valószínűsége) = 0.2.



18. ábra. Mediánszűrés eredménye. Ablakhossz = 5.

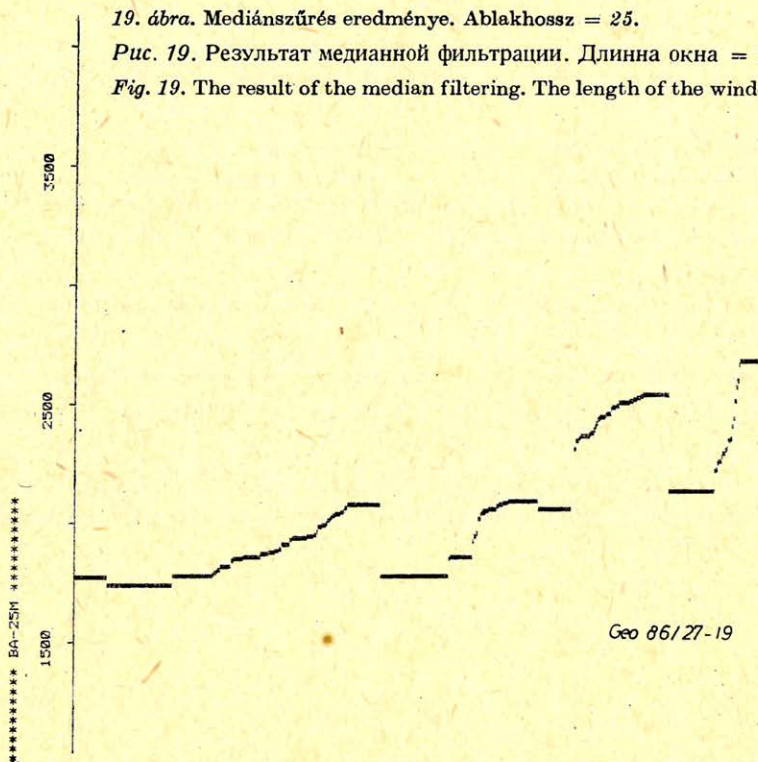
Рис. 18. Результат медианной фильтрации. Длина окна = 5.

Fig. 18. The result of the median filtering. The length of the window = 5.

19. ábra. Mediánszűrés eredménye. Ablakhossz = 25.

Рис. 19. Результат медианной фильтрации. Длина окна = 25.

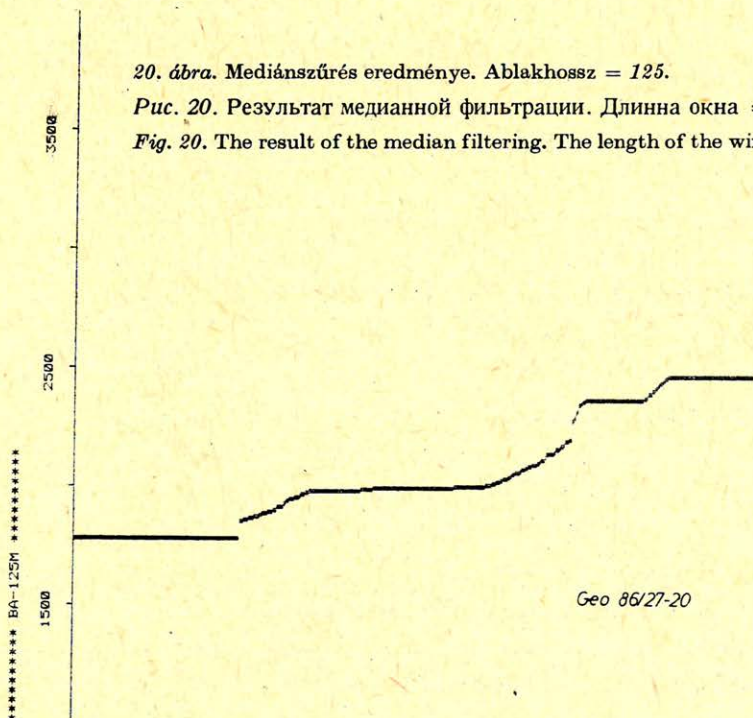
Fig. 19. The result of the median filtering. The length of the window = 25.



20. ábra. Mediánszűrés eredménye. Ablakhossz = 125.

Рис. 20. Результат медианной фильтрации. Длина окна = 125.

Fig. 20. The result of the median filtering. The length of the window = 125.



Hasonlítsuk ezt össze a 22. ábrával. Az egyedüli változás, hogy $\lambda = 0.99$. Ez a hosszú intervallumok súlyát jelentősen megnövelte. Természetesen ez a bemeneti érték követésének pontossága rovására történhetett.

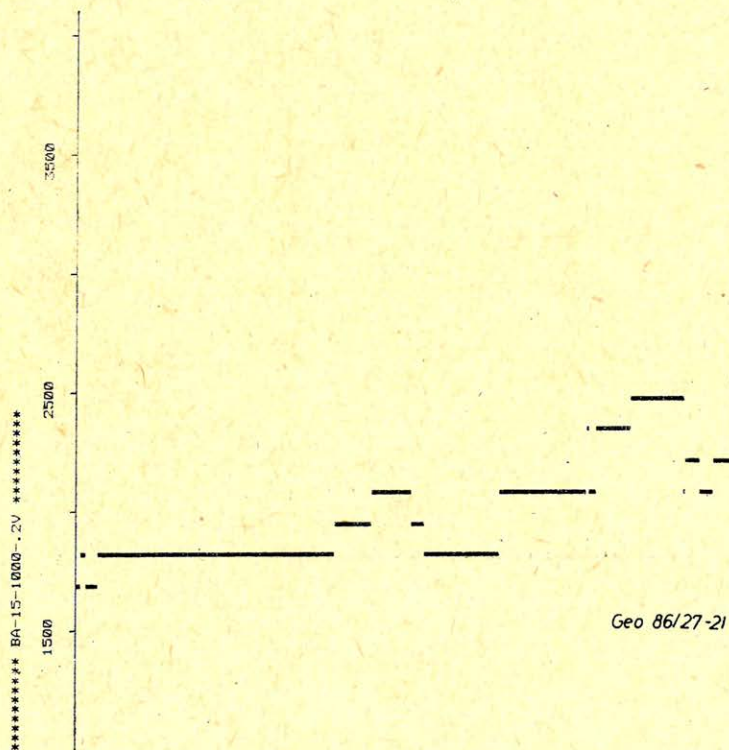
A 23. ábrán $\lambda = 0.99$ és $\sigma = 100$. A 22. ábrán ábrázolt kimenethez képest nagyobb hibajárulékot eredményez a megkívánt értéktől való eltávolodás.

A 24. ábrán σ értéke 10, így már egyedül a vélt paramétertől való legkisebb eltérésre törekvés dominál.

A paraméterek megfelelő választásával befolyásolni tudjuk a kimenet képét. A szubjektivitás az átalakításban elkerülhetetlen, hiszen – nem kimondott – célunk a bemenet információtartalmának csökkentése oly módon, hogy a kimenet könnyebben interpretálható legyen.

A 25. ábrán $\sigma = 10$, $\lambda = 0.99$ és a jelek minimuma és maximuma által határolt intervallumban 31 értéket vettünk fel.

Megállapíthatjuk, hogy a Markov-lánccokat ígéretesen használhatjuk a szegmentálási feladathoz, azonban célszerű alkalmazásához több tapasztalatra van szükség. A 25. ábra alapján is úgy tűnik, hogy érdemes kísérletezni két vagy több-dimenziós Markov-lánccokkal.



21. ábra. Markov-láncon alapuló szegmentálás eredménye. $M = 15$, $\sigma = 1000$, $\lambda = 0,2$.

Рис. 21. Результат сегментации основанной на последовательности Маркова.
 $M = 15$, $\sigma = 1000$, $\lambda = 0,2$.

Fig. 21. The result of the segmentation based upon the Markov chain model. $M = 15$, $\sigma = 1000$,
 $\lambda = 0.2$.

Összefoglalás

Áttekintettünk néhány módszert, melyek geofizikai adatrendszerek szegmentálásában a siker reményében használhatók. Különös figyelmet szenteltünk a Walsh-függvények rendszerében való előállításokra. Megvizsgáltunk átlagtól való eltérés alapján történő kvantálást, a mediánszűrésen és a Markov-láncon alapuló szegmentálást.

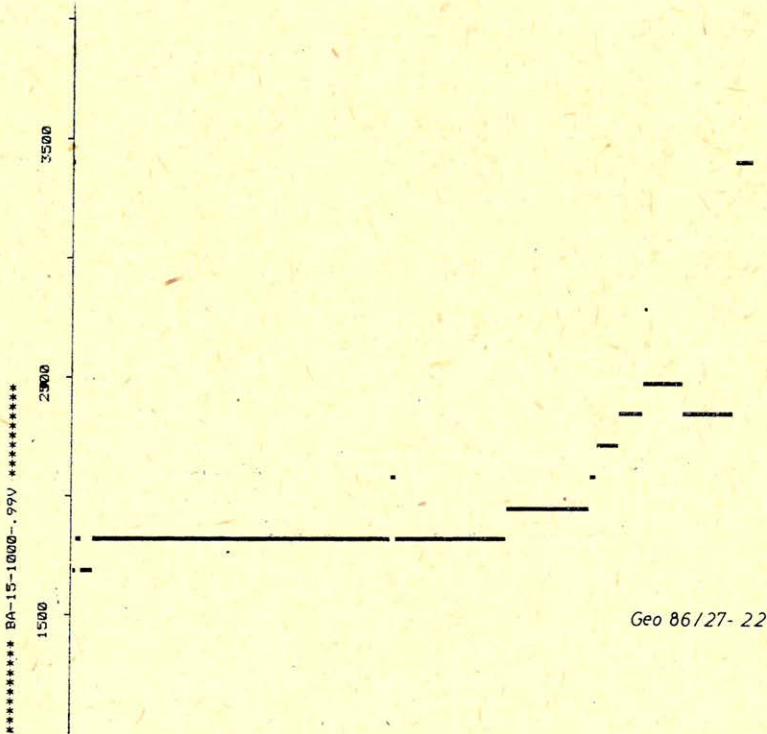
A vizsgált eljárásokat valódi adatokon, battonyai mélyfúrás akusztikus karotázsadataiból származtatott sebességadatokon próbáltuk ki.

A szegmentálási eljárás információcsökkentő hatású. Ez nem okoz problémát, ha a kapott eredmények könnyebben interpretálhatók.

Az átalakítások, hatásuk szerint, két jellegzetes csoportba sorolhatók:

Az egyiknél a lehetséges intervallumok határai kötöttek, az intervallum belsőjében tetszőleges értéket vehetnek fel. Tipikus példa a Walsh-tartományban végzett aluláteresztő szűrés.

A másik csoportban a kimeneten felvehető értékek száma korlátozott, az intervallumok határai az eljárásból adódnak. Jellegzetes példa a Markov-láncon alapuló szegmentálása.



22. ábra. Markov-láncon alapuló szegmentálás eredménye. $M = 15$, $\sigma = 1000$, $\lambda = 0,99$.

Рис. 22. Результат сегментации основанной на последовательности Маркова.

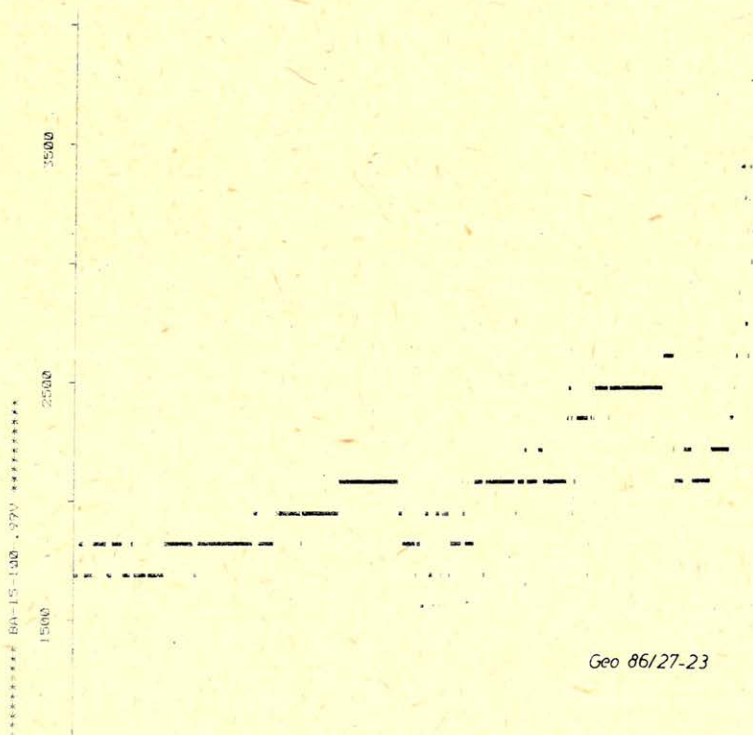
$M = 15$, $\sigma = 1000$, $\lambda = 0,99$.

Fig. 22. The result of the segmentation based upon the Markov chain model. $M = 15$, $\sigma = 1000$, $\lambda = 0.99$.

Érdeemes lenne olyan összetett módszert kipróbálni, mely a Walsh transzformált segítségével határoz meg valószínűsíthető amplitúdókat, melyeket azután a Markov-láncon alapuló eljárás használna fel. Ez utóbbi az intervallumok határainak rugalmas kijelölésével járulna hozzá a kimenet jobb értelmezhetőségéhez. Ha a szakaszonként konstans értékű kimenet két intervallumának amplitúdója csak kicsit különbözik, úgy érezzük, hogy felesleges két szakaszt kijelölni, elegendő egy intervallumot. Ha az egyik intervallum amplitúdóban lényegesen eltér a szomszédjától, de csak kevés mintára terjed ki, ugyancsak feleslegesnek érezzük lépcsőnek kijelölését.

Ezek alapján úgy fogalmazhatunk: jónak érzünk egy *lépcsősítő* eljárást, ha az egyes intervallumok szélessége nem kisebb egy alkalmas korlátnál és a szomszédos intervallumok amplitúdói jobban különböznek egymástól egy adott értéknél.

A szubjektivitást már a módszer kiválasztása magában hordozza. Ez nem feltétlenül hátrány, sőt érdekünkben áll az általunk jónak tartott modellhez közeli eredményt adó transzformáció használata. Az objektivitással kapcsolatos követelmény ott jelentkezik, hogy az átalakítást az egész adatsoron ugyanazon szabály szerint végezzük.



23. ábra. Markov-láncon alapuló szegmentálás eredménye. $M = 15$, $\sigma = 100$, $\lambda = 0,99$.

Рис. 23. Результат сегментации основанной на последовательности Маркова.
 $M = 15$, $\sigma = 100$, $\lambda = 0,99$.

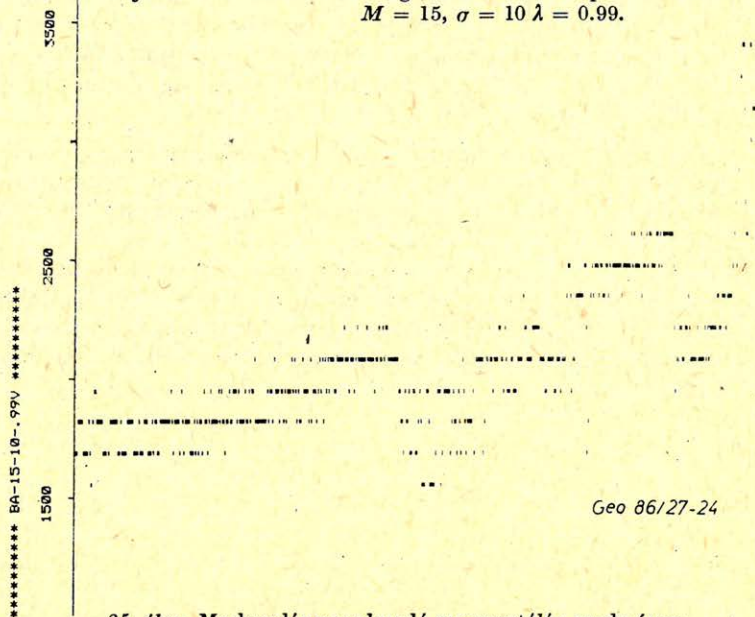
Fig. 23. The result of the segmentation based upon the Markov chain model. $M = 15$, $\sigma = 100$,
 $\lambda = 0.99$.

24. ábra. Markov-láncon alapuló szegmentálás eredménye.

$$M = 15, \sigma = 10, \lambda = 0,99.$$

Рис. 24. Результат сегментации основанной на последовательности Маркова. $M = 15, \sigma = 10, \lambda = 0,99.$

Fig. 24. The result of the segmentation based upon the Markov chain model. $M = 15, \sigma = 10 \lambda = 0.99.$

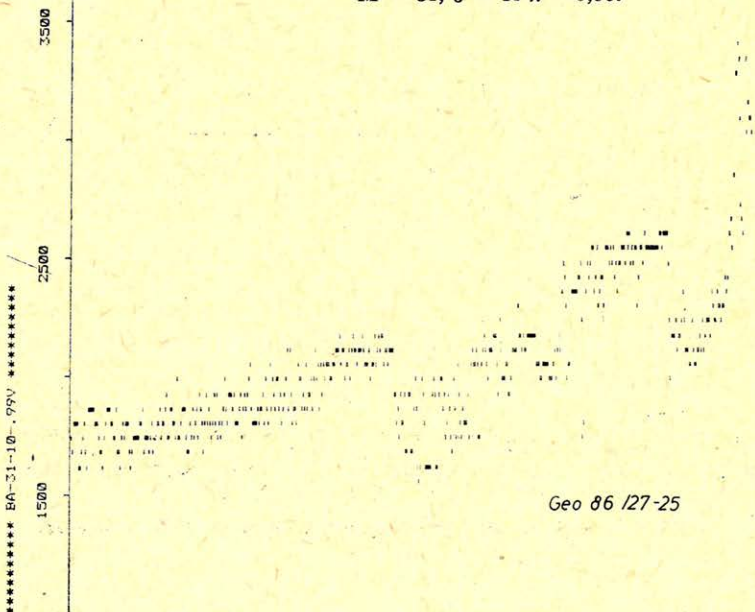


25. ábra. Markov-láncon alapuló szegmentálás eredménye.

$$M = 31, \sigma = 10, \lambda = 0,99.$$

Рис. 25. Результат сегментации основанной на последовательности Маркова $M = 31, \sigma = 10, \lambda = 0,99.$

Fig. 25. The result of the segmentation based upon the Markov chain model. $M = 31, \sigma = 10 \lambda = 0.99.$



- Beauchamp, R. G., 1975: Walsh function and their applications Academic Press.*
Davies, B., 1983: Integráltranszformációk és alkalmazásai. Műszaki Könyvkiadó.
Elliot, D. F. – K. R. Rao, 1982: Fast transforms. Algorithms, Analyses, Applications. Academic Press.
Harmuth, H. F., 1970: Transmission of information by orthogonal functions. Springer Verlag.
Kolmogorov, A. N. – Sz. V. Fomin: A függvényelmélet és a funkcionálanalízis elemei.
Meskó A., 1984: Digital Filtering. Applications in Geophysical Exploration for Oil. Akadémiai Kiadó.
Pál L. Gy., 1979: Ortogonális függvények. Tankönyvkiadó – Egyetemi jegyzet, ELTE TTK Szőkefalvi-Nagy B., 1954: Valós függvények és függvények. Tankönyvkiadó.
Vermes M., 1984: Rétegsormeghatározás karotázsszelvények számítógépes feldolgozásával. Magyar Geofizika XXV/5 – 6.

Szerkesztőségi hírek

Tagtársaink szíves tudomására hozzuk, hogy a Magyar Geofizika c. szaklap ismét megjelentet hirdetéseket.

A megjelentetni kívánt hirdetések lehetnek:

- termék, eszköz vagy gyártmányismertetés, fekete-fehér fotó vagy tusrajz illusztrációval
- tevékenységismertetés szakcikk formájában is
- szolgáltatás, know-how átadás ajánlat

A hirdetés ára:

- belső oldalanként min. 10 000 Ft
- külső oldalanként min. 15 000 Ft

A hirdetés anyagát, a megrendelő levéllel együtt az alábbi címre lehet küldeni:

MGE Magyar Geofizika Szerkesztősége

Zelei András felelős szerkesztő

Budapest Pf. 240. 1368.