

Függvénykapcsolatok megbízhatóságának növelése a leggyakoribb érték szerinti súlyok vizsgálata alapján

S Z Ü C S P É T E R *

A dolgozat a leggyakoribb érték elve során automatikusan képződő súlyparaméterek egy felhasználási lehetőségét mutatja be mélyfúrási geofizikai adatrendszer példáján keresztül.

This paper presents a possible choice of the weights by an example of a given log data system. These weights are calculated automatically by a fitting procedure according to the most frequent values.

Работа показывает возможность употребления весов по причеру системы данных ГИС. Эти веса считаются автоматически методом уравнивания наиболее частных значений.

Bevezetés

A mélyfúrási geofizikában a szelvényértelmezésben felhasznált adatsorozatok (mag- és szelvényadatok) közötti kapcsolatok vizsgálatára eddig általában hagyományos matematikai statisztikai módszereket alkalmaztak. Ezek alapvető hátránya, hogy nem robusztusak és nem rezisztensek. Azaz, csak abban az esetben működnek megfelelően és hatékonyan, ha adataink tökéletes Gauss-eloszlásúak. A gyakorlat azonban bebizonyította, hogy adataink a legkritikább esetben követik a tökéletes Gauss-eloszlást, s általában számítani kell durva hibájú, kieső adatokra is. A karotázs értelmezésben ezért mind nagyobb szerepet kell, hogy kapjon a robusztus statisztikai módszerek használata, melyek közül a dr. Steiner Ferenc által kifejlesztett leggyakoribb érték szerinti algoritmus használatát javasoljuk [1].

A szelvényértelmezésben felhasznált adatsorozatok közötti kapcsolatok vizsgálata többváltozós regressziós analízis alapján lehetséges [2]. A dolgozatban a hagyományos, legkisebb négyzetes és a leggyakoribb érték szerinti kiegyenlítési eljárások összevetése mellett, gyakorlati adatrendszeren keresztül láthatjuk a leggyakoribb érték algoritmusának olyan előnyeit, mely végül is pontossági kritériumok alapján a vizsgált változók közötti kapcsolatok megbízhatóságának növelését eredményezheti.

A hagyományos (legkisebb négyzetek elve szerinti) és a leggyakoribb értékek elve szerinti többváltozós kiegyenlítési eljárások elvi alapjai

Jelöljük $T(\bar{p}, \bar{x})$ -nal azt az ismert analitikus alakú függvényt, amelyben szereplő \bar{p} paramétervektort, azaz p_1, p_2, \dots, p_J paraméterek konkrét értékeit akarjuk valamilyen elvből következő kiegyenlítési eljárással meghatározni az (y_i, \bar{x}_i) adat-

rendszer alapján. Az adatrendszerünk n különböző pontban áll rendelkezésre. Eltérésnek nevezzük a mért y_i és a (valamilyen \bar{p} -sal) számított $T(\bar{p}, \bar{x}_i)$ különbségét.

Legyen ezek után a többváltozós kiegyenlítő függvényünk analitikus alakja a következő:

$$T(\bar{p}, \bar{x}_i) = p_1 \cdot T_1(\bar{x}_i) + p_2 \cdot T_2(\bar{x}_i) + \dots + p_J \cdot T_J(\bar{x}_i), \quad (1)$$

ahol T_1, T_2, \dots, T_J ismert analitikus alakú függvények.

Legkisebb négyzetek elve szerinti többváltozós kiegyenlítés

Ennél a módszernél az alábbi minimalizációs feltételnek kell eleget tenni:

$$\sum_{i=1}^n [y_i - T(\bar{p}, \bar{x}_i)]^2 = \min. \quad (2)$$

Ez a feltétel p_J szerinti parciális differenciálás után visszavezethető lineáris algebrai egyenletrendszer megoldására. Jól látható, hogy ennél az eljárásnál minden adatvektor egyforma súllyal szerepel. Ez a módszer tehát nem rezisztens s ráadásul nem is robusztus.

Leggyakoribb érték szerinti többváltozós kiegyenlítés

Induljunk ki a következő feltételből:

$$\sum_{i=1}^n \frac{\varepsilon^2}{\varepsilon^2 + (y_i - T(\bar{p}_k, \bar{x}_i))^2} \cdot [y_i - T(\bar{p}_k, \bar{x}_i)]^2 = \min \quad (3)$$

Ez a kifejezés mindössze a leggyakoribb érték szerinti súllyal különbözik a legkisebb négyzetek szerinti kiegyenlítés (2) feltételétől. Az ε skála-paraméter az adatvektorok leggyakoribb érték körüli tömörülésének mértékét fejezi ki. A \bar{p} és ε számítása egy kettős iterációs algoritmus alkalmazásával lehetséges. A leggyakoribb érték szerint

* NME Geofizikai Tanszék, Miskolc

kiegyenlítés egy-egy iterációs lépése csak lineáris algebrai egyenletrendszer megoldását igényli szintén, melyet a (3) p_J szerinti differenciálásával kapunk:

$$\begin{aligned} & \sum_{i=1}^n \frac{\varepsilon^2}{\varepsilon^2 + [y_i - T(\bar{p}_{k-1}, \bar{x}_i)]^2} \times \\ & \times T_1(\bar{x}_i)[p_{k1}T_1(\bar{x}_i) + \dots + p_{kJ}T_J(\bar{x}_i) - y_i] = 0 \\ & \sum_{i=1}^n \frac{\varepsilon^2}{\varepsilon^2 + [y_i - T(\bar{p}_{k-1}, \bar{x}_i)]^2} \times \\ & \times T_2(\bar{x}_i)[p_{k1}T_1(\bar{x}_i) + \dots + p_{kJ}T_J(\bar{x}_i) - y_i] = 0 \\ & \vdots \\ & \sum_{i=1}^n \frac{\varepsilon^2}{\varepsilon^2 + [y_i - T(\bar{p}_{k-1}, \bar{x}_i)]^2} \times \\ & \times T_J(\bar{x}_i)[p_{k1}T_1(\bar{x}_i) + \dots + p_{kJ}T_J(\bar{x}_i) - y_i] = 0 \quad (4) \end{aligned}$$

A súlyfüggvény skálaparaméterét szintén iterációs formulával nyerjük:

$$\varepsilon_r^2 = \frac{3 \cdot \sum_{i=1}^n \frac{\varepsilon_{r-1}^4 [y_i - T(\bar{p}, \bar{x}_i)]^2}{\{\varepsilon_{r-1}^2 + [y_i - T(\bar{p}, \bar{x}_i)]^2\}^2}}{\sum_{i=1}^n \frac{\varepsilon_{r-1}^4}{\{\varepsilon_{r-1}^2 + [y_i - T(\bar{p}, \bar{x}_i)]^2\}^2}} \quad (5)$$

ahol \bar{p} az utolsó iterációs eredményeként ismert paramétervektor.

A leggyakoribb érték szerinti kiegyenlítés algoritmusának alapvető paraméterei az egyes adatvektorokhoz tartozó súlyok, melyek értékei a leggyakoribb érték szerinti többváltozós felület és az illető adat közötti eltérés nagyságától és az ε -tól függenek:

$$q_i = \frac{\varepsilon^2}{\varepsilon^2 + [y_i - T(\bar{p}, \bar{x}_i)]^2} \quad (6)$$

Nyilvánvaló, hogy a kieső pontok, melyek az adatok zömétől távol esnek, kis súlyt fognak kapni, míg a tömörödés helyénél lévők 1-hez közelít. Jól látható (3)-ból, hogy az adatvektorok a súlyok értékének megfelelően vesznek részt a kiegyenlítésben. (Természetesen a hagyományos eljárásoknál minden adat 1-es *súllyal* szerepel.) A kieső adatvektorokat (kis súlyúak) alapvetően két fő tényező okozza:

- a mérések során adódó durva hibák;
- az adatok zömétől eltérő *viselkedés*, mely visszavezethető jelen vizsgálatunkban közetfizikai, ősföldrajzi stb. okokra.

Egyszerű, kétváltozós esetekben, amikor a pontok egy síkon ábrázolhatók, a kieső pontok

vizuálisan-manuálisan elhagyhatók, de többváltozós esetben erre már nincs lehetőségünk. A kiegyenlítés után az egyes adatvektorok súlyainak értékét ismerve megállapítható, hogy a kiegyenlítés mely vektorokat veszi kevésbé figyelembe, és esetenként mi okozza azokat.

A súlyok vizsgálatával tehát lehetőség van arra, hogy az adatok közötti, a regressziós analízis által számolt függvénykapcsolatok megbízhatóságát növeljük: például a kis súlyú adatvektorok elhagyásával, vagy az adatvektorok megfelelő csoportosításával (ha az adatrendszer pl. különböző geológiai környezetből származik).

Mivel a mélyfúrési geofizikában a különböző adatsorozatok között igen gyakran lineáris, vagy kvadratikus kapcsolatokat tételeznek fel, ezért konkrét példánkban ilyen jellegű kapcsolatok megbízhatóság-növelésének lehetőségeit mutatjuk be a leggyakoribb érték szerinti súlyok alapján [3]. A számítógépi program (C-64; IBM-AT) funkciója, hogy a lineáris vagy kvadratikus függvénykapcsolat együtthatóit meghatározza a legkisebb négyzetek módszerével és a leggyakoribb érték szerinti kiegyenlítéssel, valamint számolja a különböző hibajellemzőket és a q_i súlyokat.

Konkrét számítások és vizsgálatok

A következőkben vizsgáljuk meg egy konkrét gyakorlati adatrendszer felhasználásával, hogy a kiegyenlítés során meghatározott függvénykapcsolatok megbízhatóságát hogyan lehetne javítani és esetleg új, elsősorban közetfizikai és földtani szempontok alapján másfajta adatválogatást végezni.

Az adatrendszer

Rendelkezésünkre áll egy olyan adathalmaz [4], amely 140 db egyenként 12 különböző adatot tartalmazó adatvektorból áll. Ezek első hat eleme adott fúrásokban felvett karotázis szelvények értékeit, míg a másik hat eleme ugyanazon fúrásokból származó magok laboratóriumi magvizsgálatokból meghatározott közetfizikai- és tároló-paramétereket tartalmazza. Az adatvektorok az algyői szénhidrogénmező 3 különböző telepét reprezentálják: az Algyő-1 telepből 44 db, az Algyő-2-ből 49 db és a Szeged-1-ből 47 db adatvektor származik.

Vizsgáljuk meg, hogy például a tároló-paraméterek közül a porozitás (FI) milyen kapcsolatban van az adatrendszer szelvényeivel, illetve az azokból leszármaztatott és az ipari gyakorlatban használt mennyiségekkel (indikátorokkal), melyek a következők:

- SP-ből számított anyag-indikátor: JSP
- természetes gamma szelvényből számított anyag-indikátor: JTG
- laterológ szelvényből (R_L) és a rétegvíz fajlagos ellenállásából (R_w) számított arányszám: $R = \lg \frac{R_L}{R_w}$

— mikroellenállás adatok-

ból számított arányszám:
$$MR = \frac{R'' - R'}{R'' + R'}$$

— neutron-gamma szelvény-
ból számított indikátor:

$$NG = \exp[I_{n-\gamma}]$$

Lineáris függvénykapcsolat feltételezése esetén a

$$FI = A_1 JSP + A_2 JTG + A_3 R + A_4 MR + A_5 NG + A_6 \quad (7)$$

összefüggést kapjuk, ahol $A_1, A_2, A_3, A_4, A_5, A_6$ a meghatározandó együtthatók.

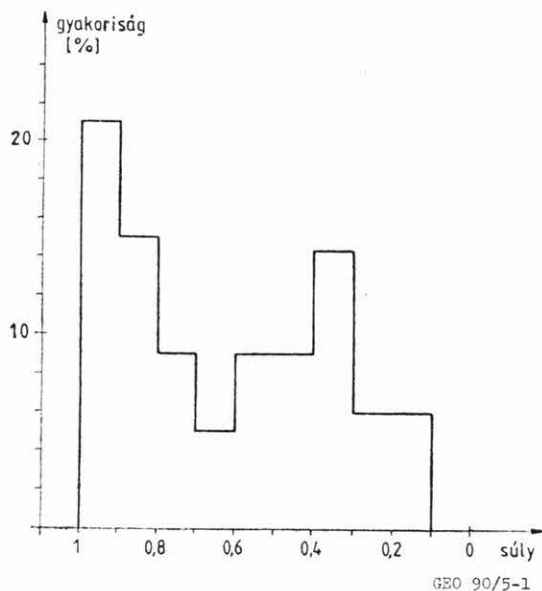
Kvadratikus esetben a kapcsolat a következő formában írható fel:

$$\begin{aligned} FI = & B_1 JSP + B_2 JTG + B_3 R + B_4 MR + B_5 NG + \\ & + B_6 JSP^2 + B_7 JTG^2 + B_8 R^2 + B_9 MR^2 + B_{10} NG^2 + \\ & + B_{11} JSP JTG + B_{12} JSP R + B_{13} JSP MR + \\ & + B_{14} JSP NG + B_{15} JTG R + B_{16} JTG MR + \\ & + B_{17} JTG NG + B_{18} R MR + B_{19} R NG + \\ & + B_{20} MR NG + B_{21}. \end{aligned} \quad (8)$$

Látható, hogy a négyzetes és keresztszorozatos tagok miatt jelentősen megnő a meghatározandó együtthatók száma: 6-ról 21-re. Emiatt a későbbiek során csak teljes adatrendszer esetén végeztünk kvadratikus kiegyenlítéseket, ugyanis az adatszám jelentősebb csökkenése az együtthatók megbízhatóságának a romlását eredményezi.

A súlyok vizsgálata

Bevezetőül szükségesnek látjuk megemlíteni, hogy példánkban elsősorban a leggyakoribb érték szerinti kiegyenlítés során képződő súlyok alkalmazási lehetőségének bemutatását tekintettük fő szempontnak, nem pedig a porozitás és a rendelkezésünkre álló szelvények közötti kapcsolat



1. ábra. Súlyok hisztogramja (teljes adatrendszer)
Рис. 1. Гистограмма весов (комплектная система данных)

Fig. 1. The histogram of the weights (full data system)

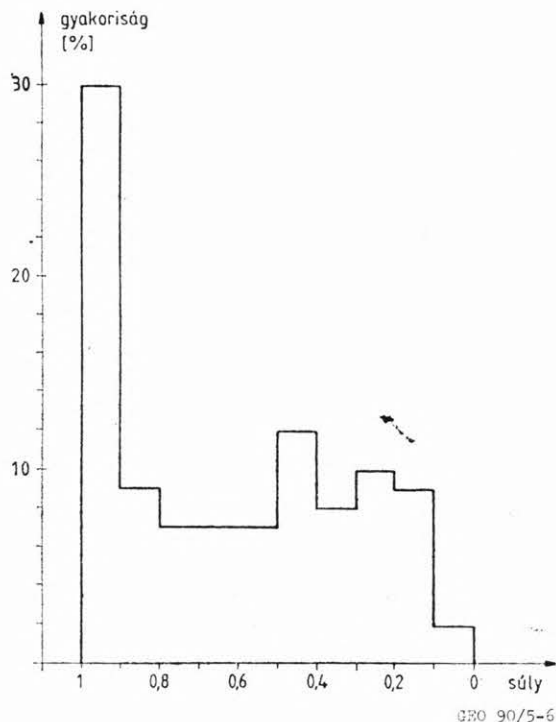
vizsgálatát. (A meghatározott függvénykapcsolatok együtthatóit ezért nem közöltük.)

Számításainkat először a teljes adatrendszer (140 adatvektor) felhasználásával kezdtük. A számítások után vizsgáltuk a statisztikus jellemzőket, valamint a súlyok hisztogramjait.

Lineáris esetre a súlyok hisztogramja az 1. ábrán található, amely a különböző súlyértékekkel rendelkező adatvektorok százalékos eloszlását mutatja. Kvadratikus esetre teljesen hasonló hisztogramot kaptunk. Vizsgálataink szempontjából a kiegyenlítés megbízhatóságáról szemléletes képet kapunk a súlyok hisztogramja alapján, ezért, valamint a könnyebb kezelhetőség miatt a későbbiekben is ezekre támaszkodunk.

Az 1. ábra súly-hisztogramján észrevehető, hogy a kis súlyoknál egy jelentős számú adatsoport található, amely a hisztogramon lokális maximummal jelentkezik. Lehetséges-e megállapítani a kis súlyú adatok viszonylag nagy gyakoriságának az okát? Ez az adatsoport valószínűleg rendelkezik valamilyen közös tulajdonsággal, ami lerontja a kiegyenlítés során nyert függvénykapcsolat szorosságát. Az okok feltárására a következőkben tárgyalandó vizsgálatokat végeztük el.

a) Végezzünk úgy kiegyenlítést, hogy elhagyjuk az eredeti adatrendszerből a kis súlyú adatok egy részét. Először a 0,2 majd a 0,3 és a 0,4-nél kisebb súlyú adatok nélkül végeztünk számításokat. Lineáris esetre a 0,4-nél kisebb súlyú adatok elhagyásával történt számítások eredményeit a 2. ábrán találhatójuk. Bár a számított



2. ábra. Súlyok hisztogramja ($q_i < 0,4$ elhagyása után)
Рис. 2. Гистограмма весов (после пропуска весов $q_i < 0,4$)

Fig. 2. The histogram of the weights (Leaving of the weights $q_i < 0,4$)

hibajellemzők javultak, a súlyok hisztogramjain továbbra is lokális maximum jelentkezett.

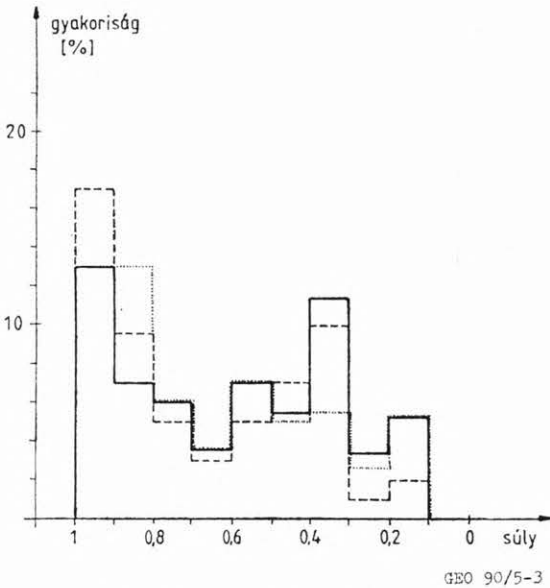
b) Lehet, hogy valamelyik szelvényérték nincs kapcsolatban a porozitással? Ennek megállapítására az 5 változós kiegyenlítést 4 változóssá változtattuk. Először a *JSP*-t, majd a *JTG*-t, *R*-t, *MR*-t és végül az *NG*-t hagytuk el a számításból.

Azt tapasztaltuk, hogy a porozitás értékének kialakításában a *JTG*-nek (*TG* szelvényből származtatott agyagindikátor) van a legnagyobb szerepe, hiszen annak elhagyásakor romlott le legjobban a közelítés. Viszont bármelyik szelvényérték elhagyásakor a kis súlyú adatok lokális maximuma továbbra is jelentkezett.

c) További számításokat végeztünk a pórus-tartalom és a permeabilitás figyelembevételével. A legtöbb adatvektor olajtároló pontokból származott. Az adatok kisebb részét képezték a víz- és gáztartalmú zónákból származó pontok. Az első számítás eredményeként kapott, az 1. ábrán látható súlyok hisztogramját úgy módosítottuk, hogy a vizes és a gázos adatokat kieső pontokként kezeltük. A lokális maximumok így is jelentkeztek.

Hasonló vizsgálatot végeztünk a permeabilitás figyelembevételével. Itt a $K \leq 50$ mD értékkel jellemezhető adatokat hagytuk el. Ez sem hozta meg a várt eredményt.

d) Mivel adataink 3 különböző telepből valók, ezért lehetőségünk van arra, hogy telepek szerint válogathassunk.



3. ábra. Súlyok hisztogramja

(— az 1. ábra Algyő—1 adatai nélkül;
 - - - az 1. ábra Algyő—2 adatai nélkül;
 az 1. ábra Szeged—1 adatai nélkül)

Рис. 3. Гистограмма весев

(— Рис. 1. без данных скважины Алдьё-1;
 - - - Рис. 1. без данных скважины Алдьё-2;
 Рис. 1. без данных скважины Сегед-1)

Fig. 3. The histogram of the weights

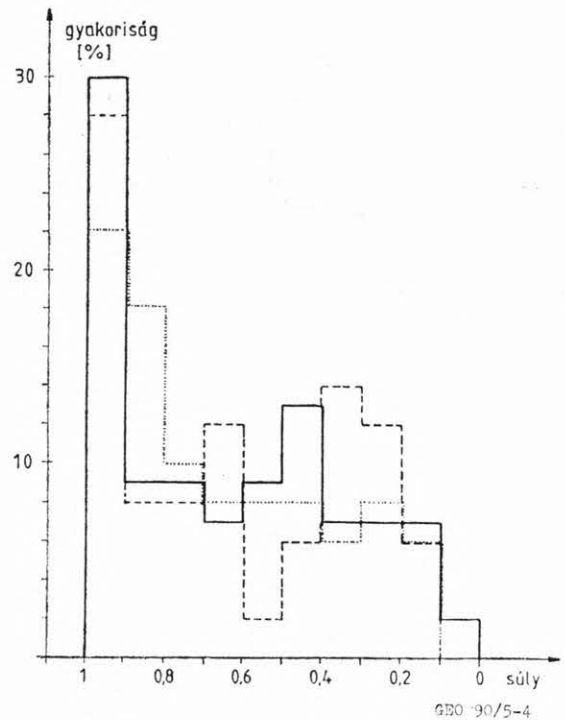
— the Fig. 1. without the data of Algyő—1;
 - - - the Fig. 1. without the data of Algyő—2;
 the Fig. 1. without the data of Szeged—1)

Az 1. ábrán lévő súlyok hisztogramját úgy alakítottuk, hogy az összes adat helyett mindig csak két telep adataiból szerkesztettük a hisztogramokat. A súlyok hisztogramjait lineáris esetben az Algyő—1, majd Algyő—2, s végül Szeged—1 telep adatai nélkül a 3. ábrán láthatjuk. Egyértelműen megállapítható, hogy a harmadik esetben, vagyis a Szeged—1 telep adatainak elhagyásával eltűnt a lokális maximum.

Annak eldöntésére, hogy valóban a Szeged—1 telep adatai okozzák-e a kisértékű súlyok lokális maximumát, további vizsgálatok szükségesek.

e) Végezzük el a számításokat mind a három telepre külön-külön. Az adatszámok így jelentősen lecsökkentek, ezért csak lineáris kiegyenlítést alkalmaztunk. A statisztikus jellemzőket az 1. táblázatba írtuk be. Azt tapasztaltuk, hogy az egyes telepekre vonatkozó kiegyenlítés során a statisztikus jellemzők elég nagymérvű javulást mutattak az eredeti kiegyenlítéshez képest. A súlyok hisztogramjai az Algyő—1, Algyő—2 és Szeged—1 telepekre vonatkozóan a 4. ábrán láthatók. Ezek alapján leghomogénebb felépítésűnek a Szeged—1 telep tűnik.

f) A külön-külön végzett számítások után vizsgálódjunk úgy, hogy két telepet összevonunk.



4. ábra. Súlyok hisztogramja

(— Algyő—1;
 - - - Algyő—2;
 Szeged—1)

Рис. 4. Гистограмма весев (— Алдьё-1;

- - - Алдьё-2;
 Сегед-1)

Fig. 4. The histogram of the weights

(— Algyő—1;
 - - - Algyő—2;
 Szeged—1)

Végezzünk további kiegyenlítéseket két-két telep adataira vonatkozóan. Nagyon érdekes, hogy Algyő—2+Szeged—1 esetre rosszabbak lettek a statisztikus hibajellemzők (1. táblázat), mint az eredeti adatrendszerre. A legjobb eredményt Algyő—1+Algyő—2 esetén kaptuk.

1. táblázat

Table 1.

Таблица 1.

Telepek szerinti kiegyenlítések statisztikus jellemzői lineáris esetben

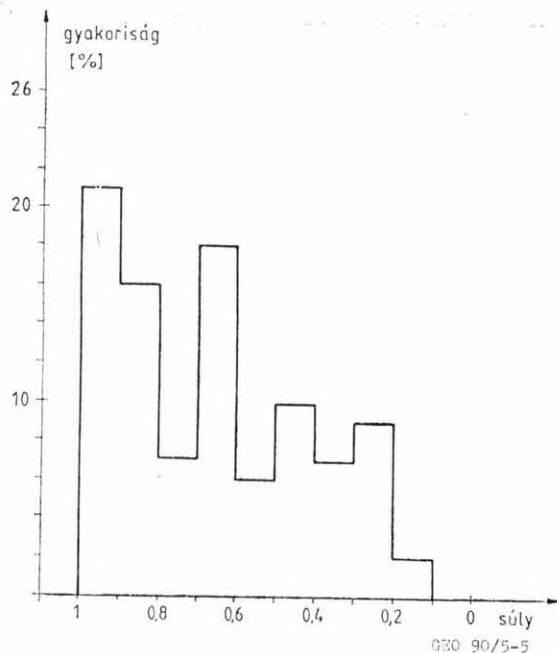
	Inter- kvartilis fél- terjedelem	Szórás	ϵ	Ki- egyen- lítés	Vek- tor- szám
Eredeti	0,01544 0,01377	0,02214 —	— 0,02086	h M	140
Algyő—I.	0,00825 0,00859	0,01601 —	— 0,01189	h M	44
Algyő—II.	0,01169 0,01194	0,02122 —	— 0,01523	h M	49
Szeged—I.	0,01724 0,01458	0,02127 —	— 0,01878	h M	47
A—II. + +Sz—I.	0,01755 0,01745	0,02373 —	— 0,02437	h M	96
A—I. + +Sz—I.	0,012921 0,01247	0,01988)—	— 0,01519	h M	91
A—I. + +A—II.	0,01307 0,01340	0,01988 —	— 0,01941	h M	93

A súlyok hisztogramjai is ezt igazolták. A legrosszabb eredményt Algyő—2+Szeged—1 telep adatai mutatták (5. ábra). A valamivel jobb kapcsolatot adó Algyő—1+Szeged—1 telep adataiból nyert súlyok hisztogramjai a 6. ábrán, míg a legjobb eredményt adó Algyő—1+Algyő—2 súlyainak hisztogramja a 7. ábrán látható. (A jobb összehasonlítás kedvéért a 7. ábrán szaggatott vonallal feltüntettük a teljes adatrendszer hisztogramját.)

Az eredmények birtokában megállapítható, hogy elsősorban a Szeged—1 telephez tartozó adatok okozzák a kis súlyok lokális maximumát.

A tárolókőzet szempontjából az egyes telep adatai következőképpen jellemezhetők [5]. Algyő—1 telep tárolókőzete agyagmárga betelepülésekkel tagolt homokkő. Apró-középszerű homokkő, finom-aprószemű homokkő és alueritos finomszerű homokkőtípusok különböztethetők meg. Algyő—2 telepnél a tárolókőzet ugyanolyan mint Algyő—1-nél, de a több agyagmárga betelepülés miatt lényegesen tagoltabb (szendvics típus), tehát az innen származó adatok változatosabbak. Ezzel magyarázható, hogy az egyes telepekre végzett kiegyenlítés során a legrosszabb eredményt az Algyő—2 telep esetén kaptuk (lásd a 4. ábrát).

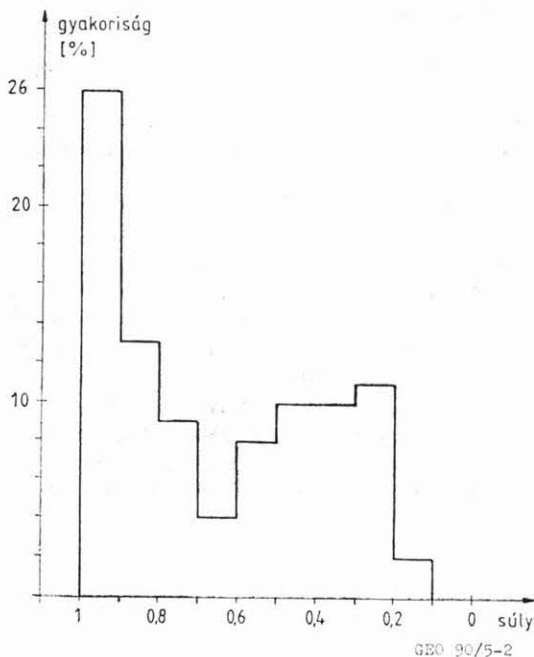
Szeged—1 telep esetében a tárolókőzet elsősorban homokkő, melyet csak ritkán tagol agyagmárga.



5. ábra. Súlyok hisztogramja (Algyő—2 és Szeged—1)

Рис. 5. Гистограмма весов (Алдьё-2 и Сегед-1)

Fig. 5. The histogram of the weights (Algyő—2 and Szeged—1)

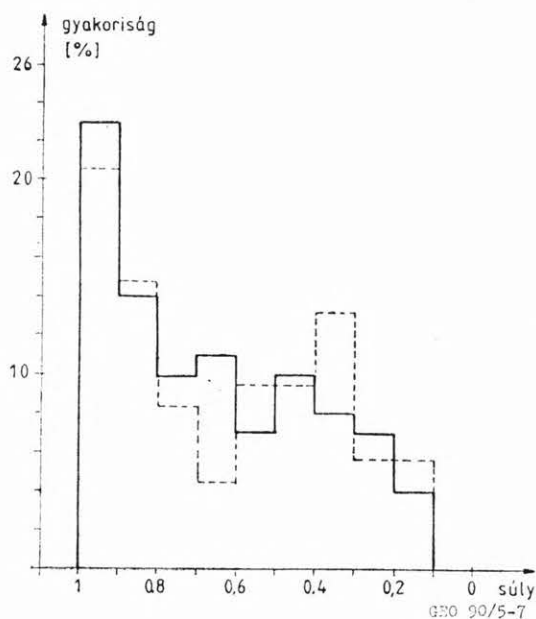


6. ábra. Súlyok hisztogramja (Algyő—1 és Szeged—1)

Рис. 6. Гистограмма весов (Алдьё-1 и Сегед-1)

Fig. 6. The histogram of the weights (Algyő—1 and Szeged—1)

A homokkő durvább szemcsés az előző két telephez képest. A telep felépítése így sokkal homogénebb az Algyő—2-nél és Algyő—1-nél is (lásd a 4. ábrát). Porozitás értékekben az egyes telepek nem nagyon különböznek (átlagosan kb. 26—27%).



7. ábra. Súlyok hisztogramja (Algyő—1 és Algyő—2)

Рис. 7. Гистограмма весев (Алдьё-1 и Алдьё-2)

Fig. 7. The histogram of the weights (Algyő—1 and Algyő—2)

Összefoglalás

A bemutatott példa is alátámasztja, hogy a leggyakoribb érték szerinti algoritmus olyan lehetőségeket teremt az alkalmazó számára, amelyet a hagyományos kiegyenlítési eljárások nem tudnak produkálni. Nevezetesen: az automatikusan képződő súlyok alapján kiszűrhetők azok az adatok (adatvektorok), amelyek a pontok tömörödési helyének „viselkedésétől” eltérnek és ezáltal a kiegyenlítés eredményeinek megbízhatóságát csökkentik. A leggyakoribb érték szerinti súlyok vizsgálata különösképpen többváltozós esetben (amikor adatainkábrázolása és így a kieső pontok vizuális érzékelésére nincs mód) adhat olyan esetleg számunkra addig ismeretlen, a geofizikai értelmezéshez szükséges lényeges információkat, amelyek alapján pl. más válogatási szisztéma, más

változók kiegyenlítésbe való bevonása, másfajta függvénykapcsolat felállítása stb. válik szükségessé.

Az előzőekben végzett számításokkal kapcsolatban a következő lényeges megállapításokat tehetjük:

a) a kiegyenlítési eljárás szempontjából

- a két kiegyenlítési eljárást összevetve a leggyakoribb értékek elve szerint nyert eredmények minden esetben megbízhatóbbaknak adódtak: a porozitás az adatok nagyobb százalékánál határozható meg ± 1 porozitátszázaléknál kisebb hibával;
- a vizsgált paraméterek között a kvadrátikus kapcsolat megfelelőbb, mint a lineáris (bár hátránya, hogy alkalmazásához — ez együtt hatók nagy száma miatt — megfelelő mennyiségű adat szükséges);

b) az értelmezés szempontjából

- a meghatározandó tárolóparaméter pontosságát jelentősen befolyásolja a megfelelő szelvények kiválasztása: jelen esetünkben például a porozitás meghatározásában legnagyobb szerepe a *JTG*-nek (*TG* szelvényből származtatott agyagindikátor) van. Elhagyása a kapcsolat szorosságát nagymértékben csökkenti;
- a leggyakoribbértékek szerinti súlyok adatvektoronkénti vizsgálata a meghatározott kapcsolatok megbízhatóságának javítását teszi lehetővé és új szempontok figyelembevételének szükségességére hívhatja fel a figyelmet;
- a különböző genetikájú (vagy földtani felépítésű) telepek együttes kezelése a kiegyenlítés hatékonyságát csökkenti, ezért elfogadható megbízhatóságú összefüggések felállítása — valószínűleg a karotázs gyakorlatban általánosan — csak telepenként végzett regressziós analízissel lehetséges.

IRODALOM

- [1] Steiner F.: Geofizikai adatfeldolgozás. Robusztus becslések. Jegyzet. Tankönyvkiadó, Budapest, 1985.
- [2] Ferenczy L.: Valószínűségelméleti alapokon nyugvó kvantitatív karotázs interpretációs rendszer hatékonyságának és megbízhatóságának növelése. Jelentés. NME. Miskolc, 1986.
- [3] Szűcs P.: A leggyakoribb érték elve szerinti többváltozós regressziós analízis alkalmazása a mélyfúrás geofizikai szelvényértelmezésben. Diplomaterv. NME. Miskolc, 1988.
- [4] Algyői magvizsgálati adatok. KV, Szolnok, 1978.
- [5] Magyarország szénhidrogén-telepei, Algyő. OKGT kiadvány, 1970.