Klaszter analízis alkalmazása felszíni geofizikai adatok komplex értelmezésében[1]

KOVÁCSVÖLGYI SÁNDOR, OCSENÁS PÉTER[2]

A rendelkezésre álló földtani–geofizikai paraméterek együttes elemzésének egyik lehetséges statisztikai módszere a klaszter analízis. A választott DK-magyarországi teszterületen a kimutatott klaszterek megfelelnek az ismert földtani szerkezeteknek, a klaszter analízis tehát alkalmas lehet e szerkezeti egységek lehatárolására, illetve az egyes egységek földtani–geofizikai sajátosságainak megismerésére.

S. KOVÁCSVÖLGYI, P. OCSENÁS: The interpretation of ground geophysical data with the application of cluster analysis

The cluster analysis is a probable statistical method of simultaneous analysis of geological–geophysical parameters. In selected area (SE Hungary) the determined clusters are the equivalent of geological structure, so the cluster analysis could be applied to determine of these structural unites and cognition of geological–geophysical quality of these unites.

1. Bevezetés

A geofizikai értelmezés módszerei determinisztikus vagy statisztikus jellegűek. Az első esetben abból indulunk ki, hogy az észlelt anomáliák világos okokra (hatókra) vezethetők vissza, a feladat pedig ezen hatók leírása. E módszerek előnye, hogy eredményük szemléletes, a földtan nyelvezetére könnyen lefordítható. Hátrányuk azonban, hogy a megbízható hatószámításhoz szükséges fizikai és geometriai paraméterek általában nem állnak rendelkezésre megfelelő mennyiségben és minőségben, ezért ezeket becsülnünk kell, illetve az egész értelmezést bizonyos — esetleg téves — prekoncepciókra kell helyeznünk. A statisztikus módszerek esetében statisztikai jellegű összefüggéseket keresünk. Ezek előnye, hogy az eredmények többékevésbé mentesek a prekoncepcióktól, földtani tartalmuk azonban nehezebben megfogható.

A különböző geofizikai módszerek a földtani felépítés különböző elemeire érzékenyek. Együttes alkalmazásuk a jobb megismerést, illetve az egyes módszerekben rejlő „csapdák" elkerülését szolgálja. A komplex értelmezés klasszikus eszközei a pauszpapír és a világítóasztal, klasszikus módszere pedig a különböző paramétertérképek szemmel történő összehasonlítása. Nyilvánvaló, hogy az összehasonlítás pontosabbá, finomabbá tehető, ha a szubjektív szem helyett objektív matematikai módszereket alkalmazunk.

Az ELGI-ben jelenleg három egymástól független, az egész országra kiterjedő adathálózat létezik: a harmadidőszaki medencealjzat mélységtérképének [KILÉNYI, SEFARA 1991] digitalizált változata, az Országos Gravimetriai Adatbázis [KOVÁCSVÖLGYI 1994, MILÁNKOVICH 1995], valamint az Országos Földmágneses Adatbázis [MILÁNKOVICH 1995]. Származtatott adathálózatként rendelkezésre áll a fedetlen gravitációs anomália térkép is. E négy paraméter klaszter analízissel történő együttes vizsgálatát kíséreltük meg a kiválasztott DK-magyarországi teszterületen.

A vizsgálat az OTKA T 019950 sz. téma támogatásával készült.

2. A klaszter analízis

A klaszter analízis feladata a mintaelemek osztályokba sorolása. A klaszterezésnek, attól függően, hogy milyen döntésfüggvénnyel végezzük az osztályozást, többféle változata ismeretes. Az eljárások két fő csoportra oszthatók: a hierarchikus és a nem hierarchikus módszerekre. Az SPSS programrendszerrel nagy elemszámú minták vizsgálatára a nem hierarchikus elemző módszerek alkalmazására van lehetőség.

A nem hierarchikus módszerek [ANDERBERG 1973] általános felépítése:

1) kezdő klaszterek kialakítása (kezdő magpontok megadása)

2) mintaelemek elhelyezése a kezdő klaszterekbe

3) a mintaelemek átrendezése a klaszterek között valamilyen optimalizáló kritérium szerint.

[1] Beérkezett: 1997. április 24-én

[2] Magyar Állami Eötvös Loránd Geofizikai Intézet, H-1145 Budapest, Kolumbusz u. 17–23.

A kezdő magpontok számításának első lépéseként tekintsük az első k adatot k középpontnak. Ezután valamennyi adatot sorban vizsgálva egy új adat minden esetben helyettesíteni fogja a hozzá legközelebb eső középpontot, ha ettől mért távolsága nagyobb, mint a két legközelebbi középpont távolsága, ill. a középpont és az új adat távolsága nagyobb, mint a középpont és a többi középpont távolsága. Az összes adaton elvégezve az elemzést rendelkezésre állnak a kezdő magpontok.

Ezután minden adatot a legközelebb lévő magponthoz sorolunk. Kiszámítjuk a klaszterek centroidjait (átlagát) és ezt tekintjük az új magpontnak. Ez utóbbi két lépést addig ismételjük, amíg a folyamat nem konvergál, azaz a klaszter tagság nem lesz állandó.

Az eljárás lényeges kérdése a klaszterek számának helyes megválasztása. Rosszul definiált klaszter számmal hamis klaszter struktúrát adhat a módszer (felbonthatunk homogén klasztereket vagy összevonhatunk elkülönülő klasztereket). A klaszterek számának helyes megválasztásához szükséges az egyes felosztásokhoz tartozó klaszterek statisztikai jellemzőinek analízise. Ez magában foglalja az egyes klaszterekhez tartozó paraméterátlagok és szórások, valamint a magpontok közötti távolságok számítását.

3. A terület földtani–geofizikai modellje

A választott tesztterület felépítése egészen az asztenoszféra felszínéig jól ismertnek tekinthető. A közelmúltban megjelent nagyszámú publikáció közül összefoglaló jellege révén kiemelkedik POSGAY et al. [1995], fejlődéstörténeti értelmezést is ad POSGAY et al. [1996], míg a gravitációs és mágneses anomáliák értelmezésével részletesebben KOVÁCSVÖLGYI [1995] foglalkozik.

Az *1. ábra* a terület pretercier aljzatának mélységtérképe, amelyen jól elkülönülnek a regionális szerkezeti elemek. Az országhatár mentén keletről nyugat felé haladva a következő szerkezeti elemek mutathatók ki:

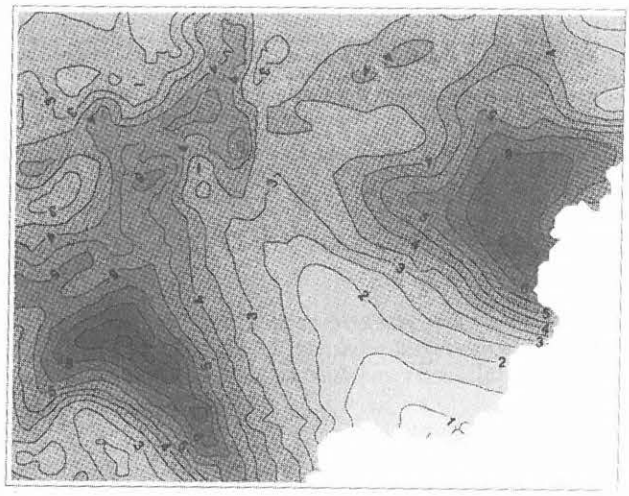
— Békési-medence, legmélyebb pontján 6 km-t meghaladó mélységgel

— Pusztaföldvár–Battonyai gerinc, legkiemelkedettebb részén 1 km körüli medencealjzattal

— Makói árok, 7 km-t elérő mélységgel

— Algyői kiemelkedés, 3 km körüli mélységgel.

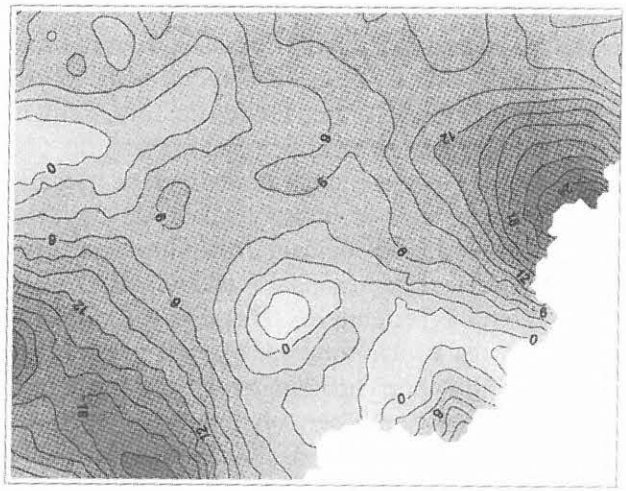
A terület ÉNy-i részén ezek a szerkezetek már nem követhetők, az aljzatmélység itt nem túl változékony, és 3–4 km körüli.



1. ábra. DK-Magyarország pretercier medencealjzatának mélységtérképe

Fig. 1. Depth to the pre-tertiary basement in SE Hungary

A Bouguer-anomália térképen (*2. ábra*) a Békési-medence és az Algyői kiemelkedés területe maximum, a Pusztaföldvár–Battonyai gerincé minimum, míg a Makói árok minimum és maximum közötti átmeneti zóna. ÉNy-on az anomália értékek a Pusztaföldvár–Battonyai gerincen mértekhez hasonlóak. A laza üledékek hatásának korrekciójával kapott fedetlen gravitációs anomália térkép (*3. ábra*) annyiban különbözik a Bouguer-anomália térképtől, hogy az Algyői kiemelkedés és a Makói árok területe egy közös maximumként jelentkezik. A felszíni mérések alapján szerkesztett mágneses anomália térképen (*4. ábra*) a Békési-medence és az Algyői kiemelkedés területén találunk regionális anomáliákat, míg ÉNy-on kisebb lokális anomáliasáv mutatható ki.

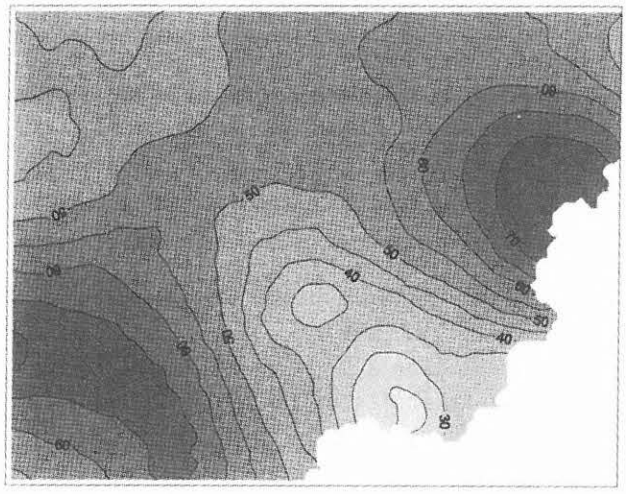


2. ábra. DK-Magyarország Bouguer-anomália (dg) térképe

Fig. 2. Bouguer anomaly map (dg) in SE of Hungary

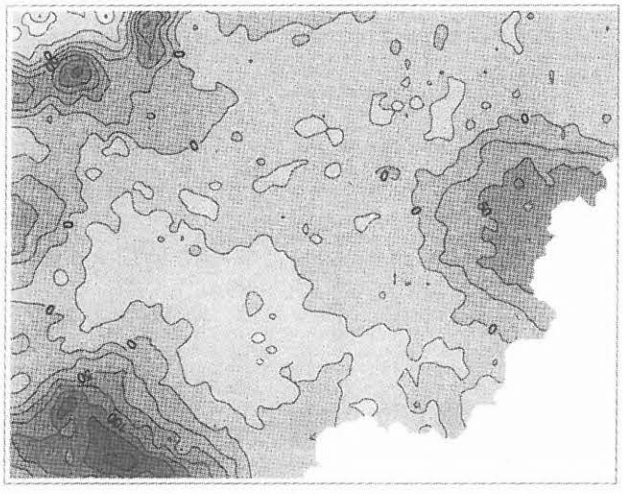
A rövid jellemzésből is kitűnik, hogy az egyes szerkezeti elemek a vizsgált négy paraméter tekin-

tetében eltérő viselkedést mutatnak, így a módszer itt megnyugtató módon tesztelhető. Ha ugyanis a klaszter analízis alkalmazásával e szerkezetek önálló klaszterekbe rendeződnek, akkor bízhatunk abban, hogy más területen, ahol az összefüggések kevésbé nyilvánvalóak, szintén értelmezhető eredményre jutunk. Ha azonban ez nem valósul meg itt, ahol a szerkezetek geofizikai sajátosságai egyértelműek, akkor a módszerrel aligha érdemes tovább kísérletezni.



3. ábra. DK-Magyarország medencehatással korrigált gravitációs anomália (fedetlen dg) térképe

Fig. 3. Stripped gravity map in SE of Hungary



4. ábra. A mágneses térerősség függőleges komponensének (dz) anomália térképe DK-Magyarországon

Fig. 4. Map of vertical magnetic component in SE of Hungary

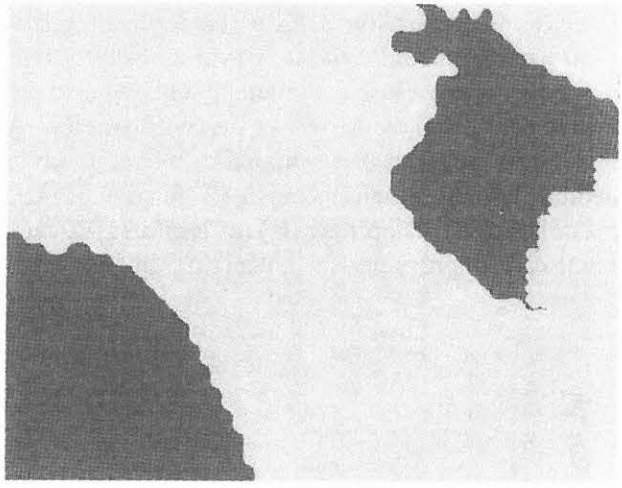
4. A klaszter analízis eredményei

A klaszter analízis első lépéseként az eredeti mérési, ill. számított adatokat 2x2 km-es négyzetrácsra interpoláltuk. Ezek lettek az analízis bemeneti adatai.

Az analízis következő lépéseként, ahhoz, hogy minden adat azonos súllyal szerepeljen, szükséges a változók standardizált értékeinek (zérus várható érték és egységnyi szórás) kiszámítása.

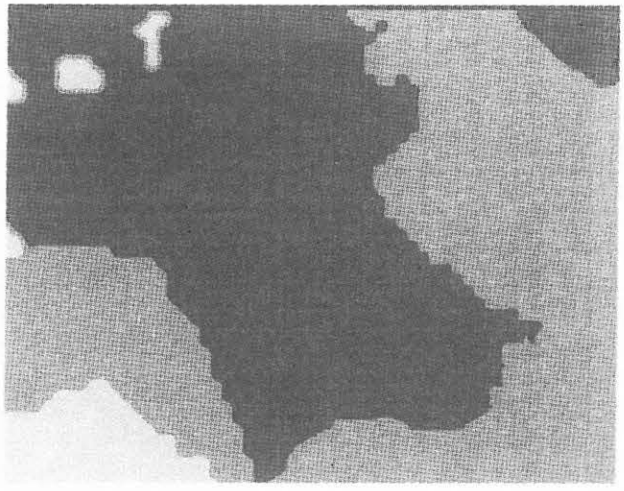
Ezt követi a konvergencia kritériumok megadása. A magpontok számítása akkor áll le, ha bármely középpont esetében a legnagyobb változás kisebb, mint a kiindulási középpontok közötti távolságnak a kritérium szerint definiált %-a. Az állandó klaszter tagság eléréséhez a tapasztalat szerint a 0,1%-os határt kell megadni. Ezzel az értékkel számolva maximum 15–20 lépés után a klaszter tagság állandósul.

Az analízist 2-, 3-, 4- és 5-klaszteres felosztásra is elvégeztük, majd az eredményeket térképen ábrázoltuk (*5.–8. ábrák*). A térképeken a különböző klaszterhez tartozás különböző árnyalattal van jelölve. Valamennyi ábra sajátossága, hogy a klaszterek területileg jól elkülönülnek.



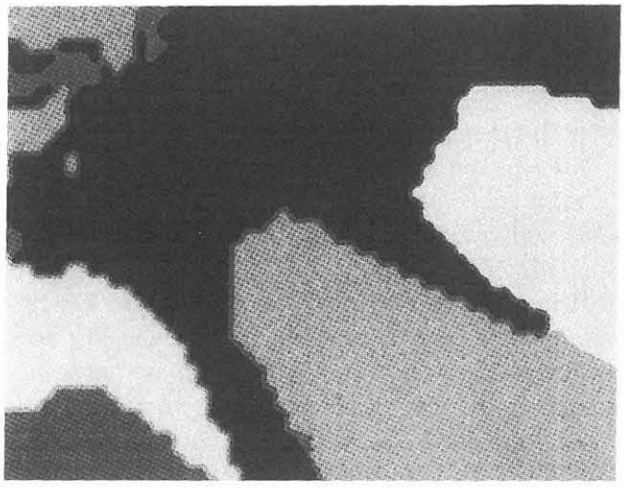
5. ábra. 2-klaszteres felosztás DK-magyarországi gravitációs és mágneses adatok alapján

Fig. 5. 2-cluster classification from gravity and magnetic data



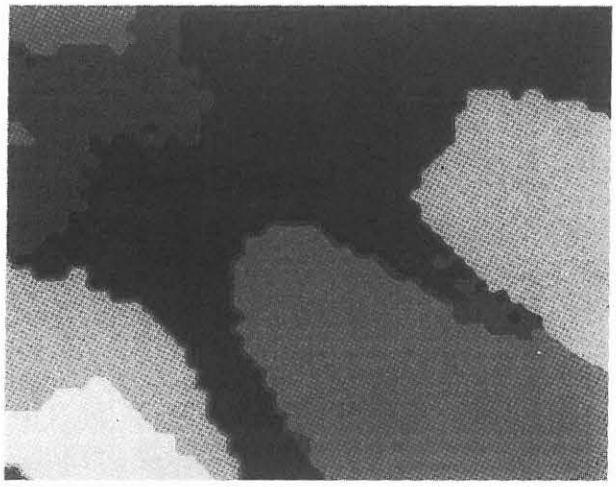
6. ábra. 3-klaszteres felosztás DK-magyarországi gravitációs és mágneses adatok alapján

Fig. 6. 3-cluster classification from gravity and magnetic data



7. ábra. 4-klaszteres felosztás DK-magyarországi gravitációs és mágneses adatok alapján

Fig. 7. 4-cluster classification from gravity and magnetic data



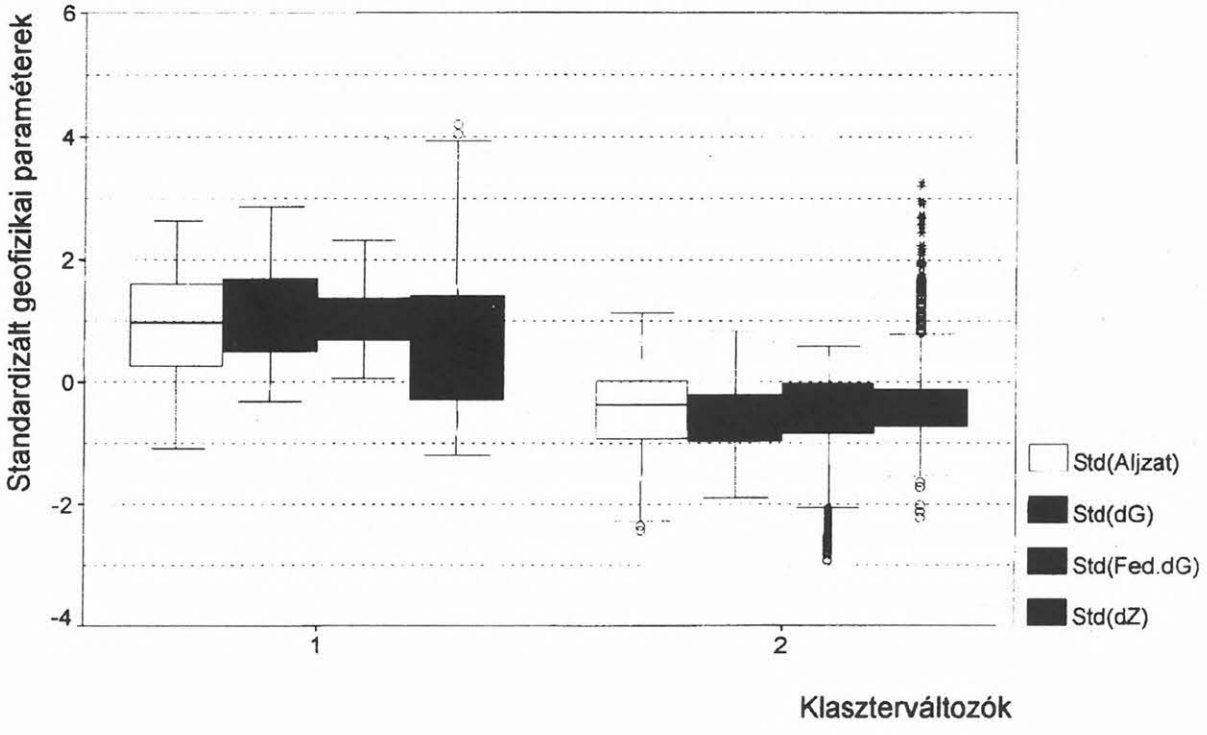
8. ábra. 5-klaszteres felosztás DK-magyarországi gravitációs és mágneses adatok alapján

Fig. 8. 5-cluster classification from gravity and magnetic data

Egyes felosztásokon belül a standardizált geofizikai paraméterek eloszlását „boxplot"-okkal szemléltetjük. Az egy eloszlást jellemző téglalap alsó éle a 25%-os, a felső éle a 75%-os percentilist jelöli. A téglalapon belül vastag vonallal a mediánt, kívül pedig folytonos vonallal az alsó, ill. felső éltől mért másfélszeres téglalap magasságot jelöltük. Az ezen kívül eső kiugró adatokat körrel (o), ill. a háromszoros téglalap magasságon túli szélsőséges adatokat csillaggal (*) jelöltük.

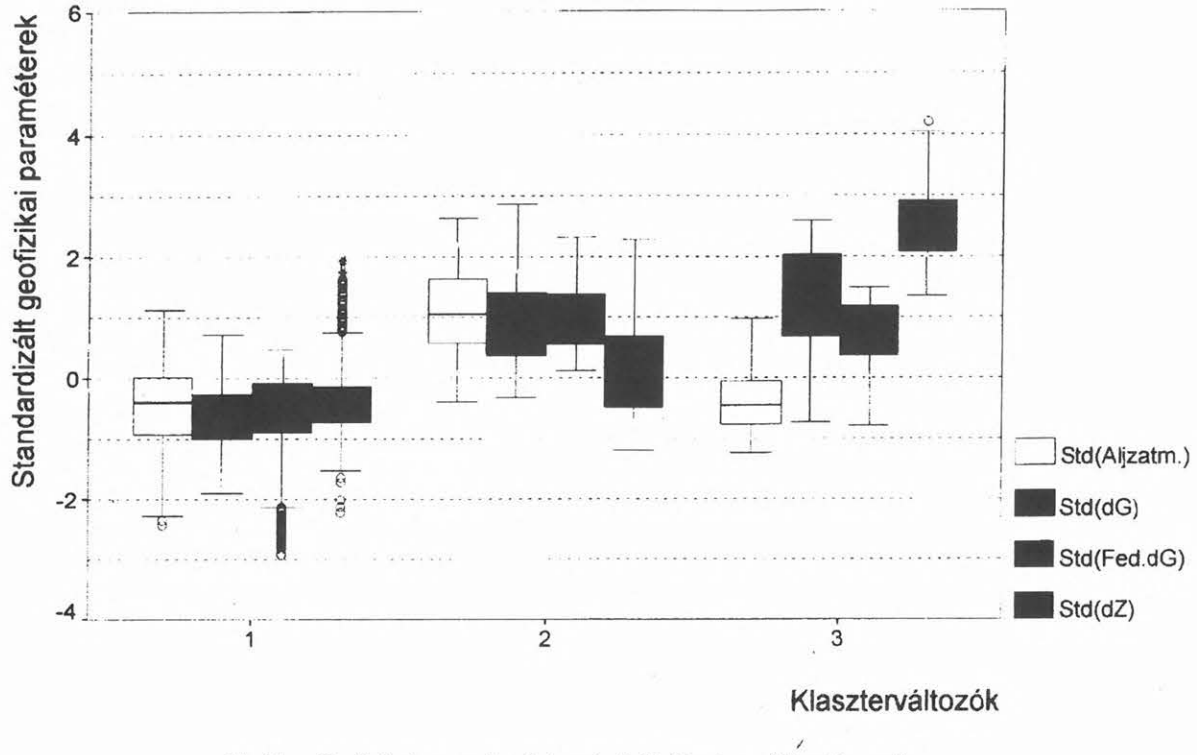
A *9.–12. ábrák*on a 2-, 3-, 4- és 5-klaszteres felosztásokhoz tartozó (standardizált) geofizikai paraméter eloszlások boxplotjai szerepelnek.

Kétklaszteres felbontás esetén (5. ábra: térkép, 9. ábra: boxplot) az első klaszterbe (az 5. ábrán sötétszürke, a 9. ábrán az 1. boxplot) a Békési-



9. ábra. Geofizikai paraméterek boxplotjai 2-klaszteres felosztás esetén

Fig. 9. The boxplots of geophysical parameters in the case of 2-cluster classification



10. ábra. Geofizikai paraméterek boxplotjai 3-klaszteres felosztás esetén
Fig. 10. The boxplots of geophysical parameters in the case of 3-cluster classification

medence, valamint a Makói árokból és az Algyői kiemelkedésből álló blokk tartozik. Itt az alsó kéreg és a köpeny kiemelt helyzetben van [POSGAY et al. 1995], ami jelentősen befolyásolja a gravitációs anomáliákat [KOVÁCSVÖLGYI 1995], olyannyira, hogy a Bouguer-anomáliák és az aljzatmélység közötti megszokott negatív korreláció helyett pozitív korrelációt kapunk. Ugyanitt a mágneses anomáliákról is feltételezhető, hogy a kiemelt helyzetű alsó kéreg képződményeihez kapcsolódnak [KOVÁCSVÖLGYI 1995]. A terület többi részén (az 5. ábrán világos színnel jelölve, a 9. ábrán a 2. boxplot) e markáns jellegzetességek hiányoznak.

A 2/1 klaszter a háromklaszteres felosztáskor (6. ábra: térkép, 10. ábra: boxplot) a 3/2 és 3/3 klaszterekre válik szét. A 3/2 klaszterbe (a 6. ábrán középszürke) kerül a Békési-medence és a Makói árok területe, míg a 3/3 klaszterben (a 6. ábrán világos) elkülönül az Algyői kiemelkedés. Az elkülönülés sajátossága az, hogy az Algyői kiemelkedés területén a nagy gravitációs anomália értékek kis aljzatmélység mellett jelentkeznek, és emellett markáns mágneses anomália is van. A kisebb aljzatmélység és a mágneses anomália együttes jelentkezése miatt ebbe a klaszterbe kerül az ÉNy-i lokális mágneses anomáliasáv területe is.

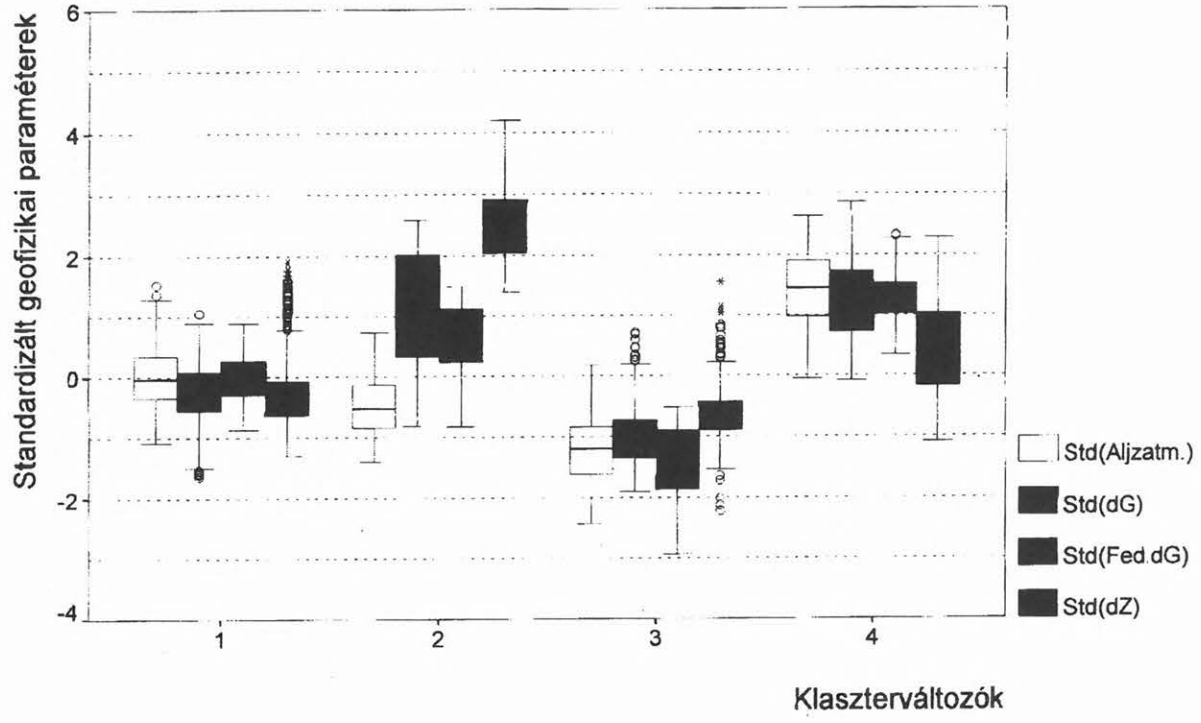
Négyklaszteres felbontásnál (7. ábra: térkép, 11. ábra: boxplot) a 4/4 klaszternek a 3/2 klaszter, a 4/2 klaszternek a 3/3 klaszter felel meg. Az új bontásban a 3/1 klaszter oszlik két részre (4/1 és 4/3). A 4/3 klaszterben a Pusztaföldvár–Battonyai gerinc területe különül el az ÉNy-i jellegtelen zónától (4/1 klaszter). A boxplot (11. ábra) mutatja, hogy a 4/3 klaszterben valamennyi vizsgált paraméter standardizált értéke negatív, míg a 4/1 klaszterban nulla körüli.

Öt klaszter esetén (8. ábra: térkép, 16. ábra: boxplot) a változás a terület ÉNy-i részét (négy klaszternél a 4/1 klasztert) érinti. Az új klaszterbe (5/2) a lokális mágneses anomália sáv és környezete kerül. A boxplot (12. ábra) szerint a továbbra is nulla körüli standardizált paraméter értékeket tartalmazó 5/1 klaszter mellett az 5/2 klaszterben a standardizált értékek általában enyhén negatívak, de pozitív mágneses anomáliákkal.

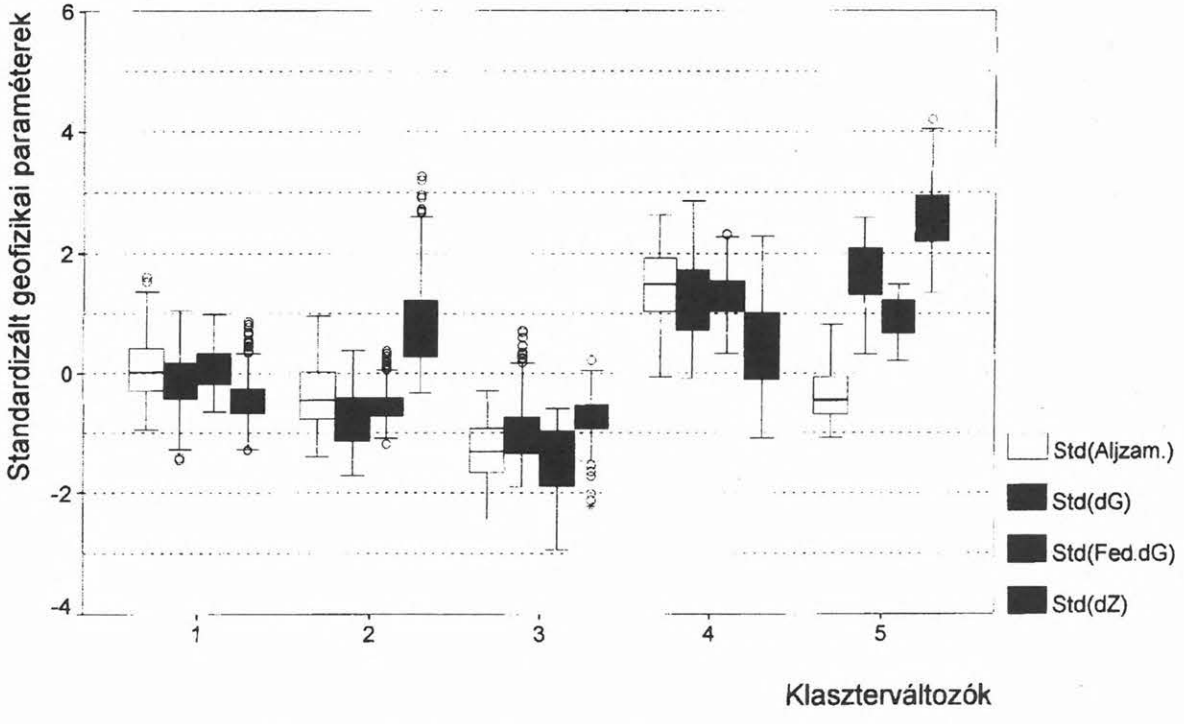
A szerkezeti egységekkel a 4-klaszteres felosztás mutatja a legtöbb hasonlóságot, így a továbbiakban ennek vizsgálatára szorítkozunk.

A 11. ábrát tovább vizsgálva, ahol az aljzatmélység, dG, fedetlen dG és dz adatok boxplotjai szerepelnek a meghatározott négy klaszterre vonatkozóan, jól látható, hogy nincs szignifikáns eltérés a dG adatok 2. és 4. klaszterekre, és a dz adatok 1. és 3. klaszterekre vonatkozó átlagértékei között.

A paraméterek páronkénti kapcsolatát szemléltetik a *13.–17. ábrák* korrelációs diagramjai. A diagramok a standardizált adatok eloszlását mutatják. Az átló menti négyzetek tartalmazzák a paraméter



11. ábra. Geofizikai paraméterek boxplotjai 4-klaszteres felosztás esetén
Fig. 11. The boxplots of geophysical parameters in the case of 4-cluster classification

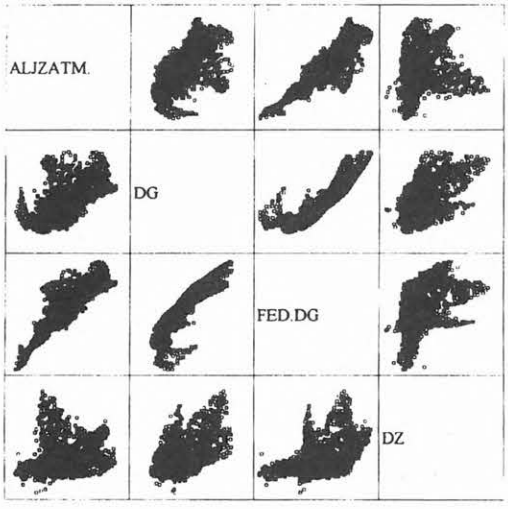


12. ábra. Geofizikai paraméterek boxplotjai 5-klaszteres felosztás esetén
Fig. 12. The boxplots of geophysical parameters in the case of 5-cluster classification

megnevezését, melynek alapján a megfelelő sor és oszlop metszésében találjuk az adott adatpárra vonatkozó diagramot.

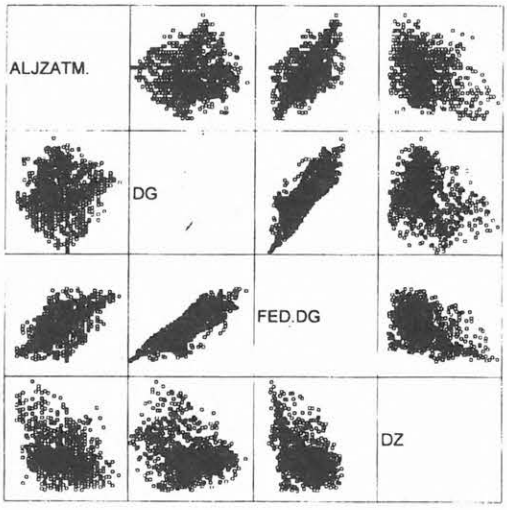
A 13. ábra a teljes adatrendszerre vonatkozik. Valamennyi adatpár esetén azt tapasztaljuk, hogy valószínűsíthetően eltérő korrelációs összefüggések

szuperponálódtak egymásra. A 14. ábra a 4/1 klaszter adatainak korrelációs viszonyait tükrözi (a 7. ábra legsötétebb területe). A terület jellegtelenségének megfelelően csak a Bouguer-anomáliák és a fedetlen gravitációs anomáliák között tapasztalunk erős korrelációt. Az aljzatmélység és a Bouguer-anomáliák a pontok egy részén korrelálnak, a többi adat esetében az ábra korrelálatlan pontfelhőket mutat. A 4/2 klaszterben (az Algyői kiemelkedés területe, 15. ábra) a korrelációs viszonyok meghatározását a kis adatszám nehezíti, a Bouguer-anomáliák és a fedetlen gravitációs anomáliák közötti jó korreláció mellett úgy tűnik, hogy a mágneses és gravitációs anomáliák között gyenge pozitív, míg a mágneses anomáliák és az aljzatmélység között gyenge negatív korreláció tapasztalható. A 4/3 klaszterben (Pusztaföldvár–Battonyai gerinc, 16. ábra) a legjobb (pozitív) korreláció az aljzatmélység és a fedetlen gravitációs anomáliák között mutatható ki. Mivel a terület mágnesesen anomáliamentes, a vizsgált paraméterek a mágneses anomáliáktól függetlenek. A 4/4 klaszterben (Békési-medence és Makói árok, 17. ábra) a mágneses anomáliák valamennyi vizsgált paraméterrel pozitív korrelációt mutatnak. Az aljzatmélység és a gravitációs anomáliák között az összefüggés összetett, úgy tűnik, hogy pozitív és negatív korreláció egyaránt jelen van.



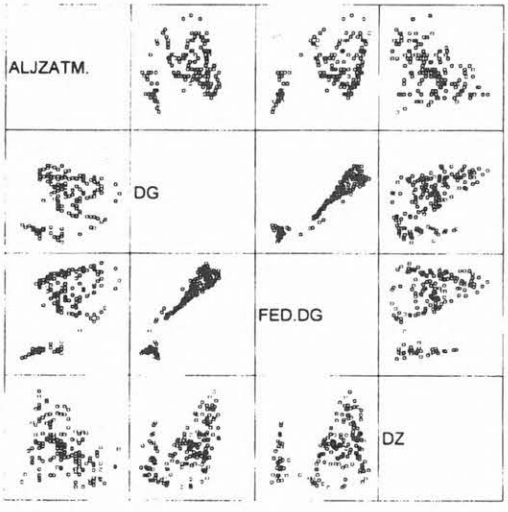
13. ábra. A geofizikai paraméterek crossplotjai (az összes adat figyelembevételével)

Fig. 13. The crossplots of geophysical parameters from all data



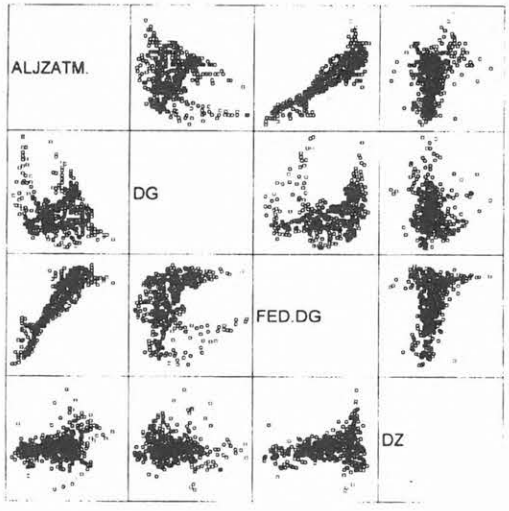
14. ábra. A geofizikai paraméterek crossplotjai (az 1. klaszter adatainak figyelembevételével)

Fig. 14. The crossplots of geophysical parameters from data of 1. cluster



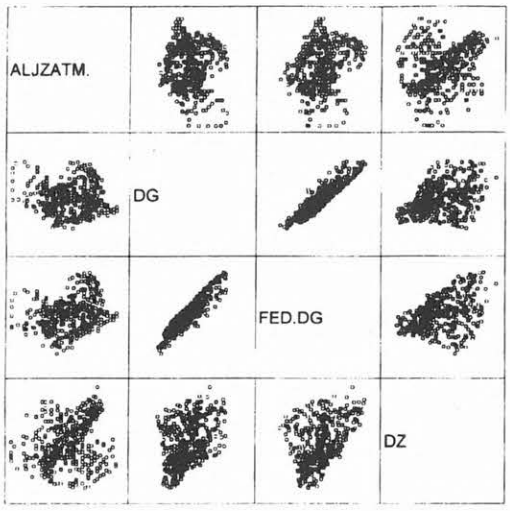
15. ábra. A geofizikai paraméterek crossplotjai (a 2. klaszter adatainak figyelembevételével)

Fig. 15. The crossplots of geophysical parameters from data of 2. cluster



16. ábra. A geofizikai paraméterek crossplotjai (a 3. klaszter adatainak figyelembevételével)

Fig. 16. The crossplots of geophysical parameters from data of 3. cluster



17. ábra. A geofizikai paraméterek crossplotjai (a 4. klaszter adatainak figyelembevételével)

Fig. 17. The crossplots of geophysical parameters from data of 4. cluster

A vizsgálati eredmények azt mutatják, hogy a klaszter analízis a geofizikai értelmezés hasznos eszköze lehet. Mint tapasztaltuk, a tesztterület klaszterei a földtani szerkezet egységeit területileg követik, az adatok klaszteren belüli elemzése pedig segítséget nyújt a szerkezeti egységek geofizikai sajátosságainak megismerésében.

HIVATKOZÁSOK

ANDERBERG H. R. 1973: Cluster analysis for application. New York, Academic Press

KILÉNYI É., ŠEFARA J. 1991: Pre-tertiary basement contour map of the Carpathian basin beneath Austria, Czechoslovakia and Hungary. Geophysical Transactions **36**, enclosure

KOVÁCSVÖLGYI S. 1994: Az ELGI gravimetriai adatbázisa. Magyar Geofizika **35**, 1, 44–47

KOVÁCSVÖLGYI S. 1995: DK-Magyarország gravitációs és földmágneses anomáliáinak értelmezése. Magyar Geofizika **36**, 3, 198–202

MILÁNKOVICH A. 1995: Jelentés a gravitációs, földmágneses, geoelektromos és mérnökgeofizikai adatbázis fejlesztéséről. MGSz Geofizikai Adattár, AD. 1430

POSGAY K., BODOKY T., HEGEDŰS E., KOVÁCSVÖLGYI S., LENKEY L., SZAFIÁN P., TAKÁCS E., TÍMÁR Z., VARGA G. 1995: Astenospheric structure beneath a Neogene basin in south-east Hungary. Tectonophysics **252,** 467–484

POSGAY K., TAKÁCS E., SZALAY I., BODOKY T., HEGEDŰS E., J. KÁNTOR I., TÍMÁR Z., VARGA G., BÉRCZI I., SZALAY A., NAGY Z., PÁPA A., HAJNAL Z., REILKOFF B., MUELLER S., ANSORGE J., DE IACO R., ASUDEH I. 1996: International deep reflection survey along the Hungarian Geotraverse. Geophysical Transactions **40,** 1–2, 1–44