MISKOLCI EGYETEM

MÓDSZERFEJLESZTÉSEK HETEROGÉN MÉRNÖKI ADATHALMAZOKHOZ

PhD ÉRTEKEZÉS

Készítette

Kilik Roland
okleveles mérnök-informatikus

HATVANY JÓZSEF INFORMATIKAI TUDOMÁNYOK DOKTORI ISKOLA
TERMELÉSINFORMATIKA TÉMATERÜLET
MÉRÉS- ÉS IRÁNYÍTÁSTECHNIKAI INFORMÁCIÓS RENDSZEREK TÉMACSOPORT

Doktori Iskola Vezető
Prof. Dr. Kovács László
egyetemi tanár, DSc

Témavezetők
Prof. Dr. Szabó Norbert Péter
egyetemi tanár, DSc
Dr. Varga Attila Károly
egyetemi docens, PhD

Miskolc
2026

NYILATKOZAT

Alulírott, Kilik Roland, kijelentem, hogy ezt a doktori értekezést magam készítettem, és abban csak a megadott forrásokat használtam fel. Minden olyan részt, amelyet szó szerint, vagy azonos tartalomban, de átfogalmazva átvettem, egyértelműen, a forrás megadásával megjelöltem.

A dolgozat bírálatai, és a védésről készült jegyzőkönyv a későbbiekben a Miskolci Egyetem Gépészmérnöki és Informatikai Karának Dékáni Hivatalában lesznek elérhetők.

Miskolc, 2026. 03.08

Kilik Roland

KÖSZÖNETNYILVÁNÍTÁS

Jelen értekezés a Hatvany József Informatikai Tudományok Doktori Iskola keretein belül, a Miskolci Egyetem Geofizikai Intézeti Tanszékén végzett kutatómunkám eredményeit foglalja össze.

Ezúton szeretném kifejezni hálámat mindazoknak, akik tanácsaikkal és támogatásukkal hozzájárultak a disszertáció megvalósulásához. Külön köszönettel tartozom témavezetőmnek, **Dr. Szabó Norbert Péter** professzor úrnak értékes szakmai útmutatásaiért, valamint a különböző projektekbe való bevonásért, ami alapvetően meghatározta kutatásom sikerét és a publikációs követelmények teljesítését.

Hálás vagyok **Dr. Dobróka Mihály** professzor úrnak, aki mindvégig figyelemmel kísérte szakmai fejlődésemet, és építő jellegű javaslataival segítette tudományos előrehaladásomat. Köszönetemet fejezem ki továbbá **Dr. Varga Attila Károly** társtémavezetőmnek, aki szakmai iránymutatásával és támogató közreműködésével nagymértékben hozzájárult a disszertáció sikeres elkészültéhez.

A tanulmány elkészítését az MTA Fenntartható Fejlődés és Technológiák Nemzeti Program (FFT NP FTA) támogatta. A kutatás részben a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal támogatásával, a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott ADVANCED-152958 azonosítószámú projekt keretében valósult meg. A disszertációban ismertetett osztályozási eljárás előzménykutatása a TÁMOP-4.2.2.C-11/1/KONV-2012-0002 azonosítószámú projekt keretében valósult meg.

TÉMAVEZETŐI NYILATKOZAT

Kilik Roland kutatásai során heterogén adatrendszerek osztályozásával, klaszterelemzésével és zajszűrésével foglalkozott, vizsgálva az informatikai és a földtudományi területek kapcsolódó szegmenseit. Különösen kiemelendő, hogy a jelölt az adatok minőségét javító, robusztus és adaptív módszereket dolgozott ki, így például: a hanginformációkon alapuló fejmozgás-osztályozást, az outlier-detektálás hatékonyságát növelő kombinált Mahalanobis- és euklideszi távolságra épülő eljárást, a hisztogram-alapú zajszűrést MFV súlyozással, valamint a heterogén geofizikai adatokra alkalmazott MFV-alapú klaszterelemzést és annak adaptív súlyozott kiterjesztését több mélyfúrás adatainak együttes feldolgozására.

Kilik Roland kutatási eredményeit hazai és nemzetközi konferenciákon és rangsorolt folyóiratokban publikálta. Az MTMT adatbázis alapján 30 közleménye jelent meg, 1 könyvrészlet szerzője. Független hivatkozásainak száma 32, ami a kutatások nemzetközi elismertségét is tükrözi.

Az értekezés multidiszciplináris kutatáson alapuló tudományos munkát tükröz. A jelölt által kidolgozott módszerek a mérnöki gyakorlat számára is jelentőséggel bírnak, mivel hozzájárulnak a heterogén mérnöki és geofizikai adatok hatékonyabb feldolgozásához, valamint az adatelemzési eredmények megbízhatóságának növeléséhez. Az értekezés logikusan felépített, és átfogó képet ad a jelölt módszerfejlesztési kutatásairól. Az alkalmazási lehetőségek többek között a nyersanyagkutatásban, hidrológiai és geofizikai adatok elemzésében, valamint a heterogén adatok korszerű feldolgozásának különböző területein jelennek meg.

A disszertáció három fő területen tartalmaz új tudományos eredményeket: (i) hanginformáció-alapú osztályozás és mozgásmérték-becslés, (ii) hisztogram-alapú zajszűrés és outlier-kezelés heterogén adatokon, (iii) MFV-alapú klaszterelemzés adaptív súlyozással, valamint a klaszterezés jóságát mérő Hurst-exponens alkalmazása.

A disszertáció Kilik Roland önálló kutatási eredményeit tartalmazza, és minden szempontból megfelel a Hatvany József Informatikai Tudományok Doktori Iskola szabályzatában előírt követelményeknek. Az értekezés jelentős tudományos és gyakorlati hozzájárulást képvisel a jel- és adatfeldolgozás területén.

A fentiek alapján a jelölt számára a PhD cím odaítélését támogatjuk.

Miskolc, 2026. március 12. Dr. Szabó Norbert Péter és Dr. Varga Attila Károly

Tartalomjegyzék

Bevezetés

Még graduális képzésem során látókörömbe került és érdeklődési területemmé vált az adatfolyamok, adatrendszerek mögöttes információinak feltárása, illetve adatok javítása. Diplomamunkámként streaming adatfolyamok valós idejű elemzését lehetővé tevő szoftvert valósítottam meg, majd saját ötletként vezeték nélküli IPTV adatfolyamok robusztussá tételének lehetőségét kutattam, és fejlesztettem ki arra többedmagammal egy megoldást. Ezt követően került látókörömbe az akkor fő tanszéki kutatási területet jelentő beszédfeldolgozás-képfeldolgozás témaköre, így az adatokhoz való affinitásom ebben a témában igyekeztem kamatoztatni, amely az értekezés első részében bemutatott osztályozási megoldás megalkotásában testesült meg. A műszaki földtudományok kutatásaiba bekapcsolódva, az ott alkalmazott adatvizsgálati, zajszűrési módszereket és azok korlátait megismerve támadt ötletem arra, hogy az ismert módszereket is felhasználva, de mégis alapjaiban új zajszűrési alternatívát alkossak, amelyet a dolgozat második pontjában ismertetek. Ezen eljárás eredményei ígéretesnek mutatkoztak arra, hogy a klaszterezés területén is megoldást próbáljak találni ennek alapjain, amelynek kutatási eredményeit a disszertáció harmadik pontjában ismertetem.

A mindennapjainkat és jövőnket is egyre inkább befolyásoló különböző rendszerek nem tudnak működni adatok nélkül, így az adatok feldolgozása, az információk kinyerése, az adathalmazok javítása talán fontosabb, mint valaha. Gondoljunk csak például arra, hogy a generatív mesterséges intelligencia megoldások is csak annyira lehetnek jók, amilyen minőségű adathalmazon tanították azokat, illetve amilyen mértében robusztusak tudnak lenni a kieső értékekkel szemben. Nem csupán a világunkban keletkező napi adatmennyiség nő folyamatosan [1], de az adatokat felhasználó – például deep learning – megoldások száma és jelentősége is [14]. Így a személyes motiváción túl az adatok előszűrésének, osztályozásának, információkinyerésének fontossága nem csak további motivációt jelentett kutatásaimhoz, de meghatározza relevanciáját is.

Az értekezés három fő területen tartalmaz új tudományos eredményeket, melyek alapvetően a jel- és adatfeldolgozás témaköréhez tartoznak. Az első egy újonnan kifejlesztett osztályozási eljárás, melyet egy hisztogramalapú súlyozott medián zajszűrő eljárás, végül pedig az ezen alapuló robusztus klaszterelemzési eljárás követ. Mindegyik eljárás alapját az adott adatrendszerek előállását követően azok manuális elemzése, heurisztikák, majd a kezdeti eredmények, észrevételek felhasználásával megalkotott szabályszerűségek adták, amelyek az iteratív módszerfejlesztések kiindulópontjai voltak. Mindhárom módszert illetően közös

kiindulópont és egyben munkahipotézis volt az adatok előszűrése olyan értelemben, hogy az osztályozási eljárásnál is kiinduló feltételezésem volt, hogy jobb eredményt produkáló módszer alkotható meg, ha elérhető mennyiségek bizonyos halmazainak adott határok közötti mintáira kerülnek meghatározásra az osztályt alkotó szabályok (kvázi előszűrést alkalmazva), és ugyanígy előszűrést alkalmaztam – más, hisztogramos formában – a zajszűrésre, majd ezt alapul vevően a klaszterezésre megalkotott módszer esetében is.

1. Osztályozási eljárás virtuális ügynökök mozgásának generálásához

1.1 Bevezetés

Az emberhez hasonló ágensek a mesterséges intelligencia előretörésével egyre gyakoribbá válnak. Ezen ágensek hasznossága azonban nagymértékben függ mozgásuk természetességétől.

Számos tanulmány kimutatta, hogy egy fejmozgást szintetizáló rendszer létrehozható különböző hanginformációt bemenetként használva, miközben kevés, a vizuális információkból kinyert referenciát vesz ehhez alapul [63]. Korábbi kutatások megerősítették, hogy a cél nem az eredetihez minél közelebbi fejmozgások előállítása abban az értelemben, hogy a mozgás természetességének mértéke nem lehet a referencia trajektóriájához való hasonlóság mértéke. Ez a megfigyelés egybeesik a különböző alanyok neurális hálóval szintetizált fejmozgásairól kapott kezdeti minőségi visszajelzésekkel. [S1] A kezdeti vizsgálat során ugyanis egyértelműen bebizonyosodott, hogy a generált fejmozgás trajektóriájának az eredeti (lekövetett) mozgás trajektóriájához való hasonlóságával szemben az eredetihez közeli maximális fejmozgás-amplitúdó eredményez természetesebb benyomást. Ez a megállapítás vezetett a fejmozgás amplitúdójának hanginformációból történő meghatározására irányuló vizsgálataimhoz, azzal a céllal, hogy a becsült amplitúdóérték felhasználható legyen neurális hálózat által generált fejmozgások automatikus amplitúdó-korrekciójához, ráadásul azzal az elvárással, hogy a folyamat során ne legyen szükség vizuális információra.

Az automatikus mozgásgenerálásról szóló, korábbi eredményeket, illetve módszerek korlátait összehasonlító fontos eredményeket közöltek Zhou és szerzőtársai [66]. Ebben a szerzők rámutatnak, hogy a virtuális beszélő fejek élethű mozgásának generálásában az egyik fő probléma az, hogy a nagyon alapvető hanginformációs mennyiségek önmagukban nem korrelálnak kellőképpen a mozgással, így a csak hang alapján történő élethű (pl. megfelelő amplitúdójú) fejmozgás generálása nagy kihívást jelentő feladat.

A fenti szerzők hivatkoznak Hyeongwoo Kim és szerzőtársai [30] munkájára is, ahol a szerzők azt vizsgálták, hogy a mozgásgenerálás minőségének nagymértékű romlását okozza, ha a becsült/generált mozgástartomány kívül esik az adott mozgásgeneráló rendszer tanítási mintáinak mozgástartományán. Ez is megerősíti a fejmozgás amplitúdójának a hangból történő becslésének – és természetesen a mozgásgenerálás során történő felhasználásának – fontosságát.

Az általam kifejlesztett osztályozási módszer célja az emberszerű ágensek fejmozgásainak természetességének növelése, mégpedig a valós fejmozgás mértékének pusztán a hanginformációkból történő becslésével, amelynek ismeretében a virtuális beszélő fejek, ill. robotok fejmozgás mértékei abba a tartományba skálázhatók, amelyben egy valós beszélő feje mozgott volna. Eljárásom segítségével automatikusan korrigálható a generált mozgás tartománya végső fázis-amplitúdó korrekció formájában a természetesség növelése érdekében. Főbb jellemzői, hogy az alanyok maximális függőleges fejmozgását csak a hangmintákból határozza meg, általános alanyok esetében is működik, pontossága és hatékonysága további osztályok meghatározásával javítható, és bármely fejmozgás generáló módszer kimenetének természetességét javíthatja egy utólagos amplitúdóskálázással. Ezen fő előnyök, illetve alkalmazási lehetőség mellett a fejmozgás-tartományok és az osztályozásban használt származtatott hanginformációs mennyiségek közötti magas korrelációs érték továbbfejlesztési lehetőségek kiindulópontja lehet [S2]. Mindemellett a módszer a későbbiekben esetlegesen alkalmas lehet annak a vizsgálatára is, hogy egy videóban szereplő beszélő valódi-e, vagy azt (annak mozgását) valamilyen generatív mesterséges intelligencia segítségével állították-e elő szintetikusan. Ennek, mint lehetséges alkalmazási lehetőségnek a jelentősége vélhetően a közeljövőben igen nagy lesz.

Módszerem megalkotását követően megvizsgáltam eljárásomat illetően az osztályszintű összefüggéseket, illetve az osztályokban található változók számát és az osztályok esetleges összevonhatóságát is. Ilyen értelemben tehát a módszer által kapott osztályszerkezeti optimalitásnak a vizsgálata is megtörtént, amely szintén bemutatásra kerül. Az értekezés tartalmaz egy újszerű outlier-detektálási eljárást is, amelynek meghatározó szerepe van az osztályozás hatékonyságának további javításában.

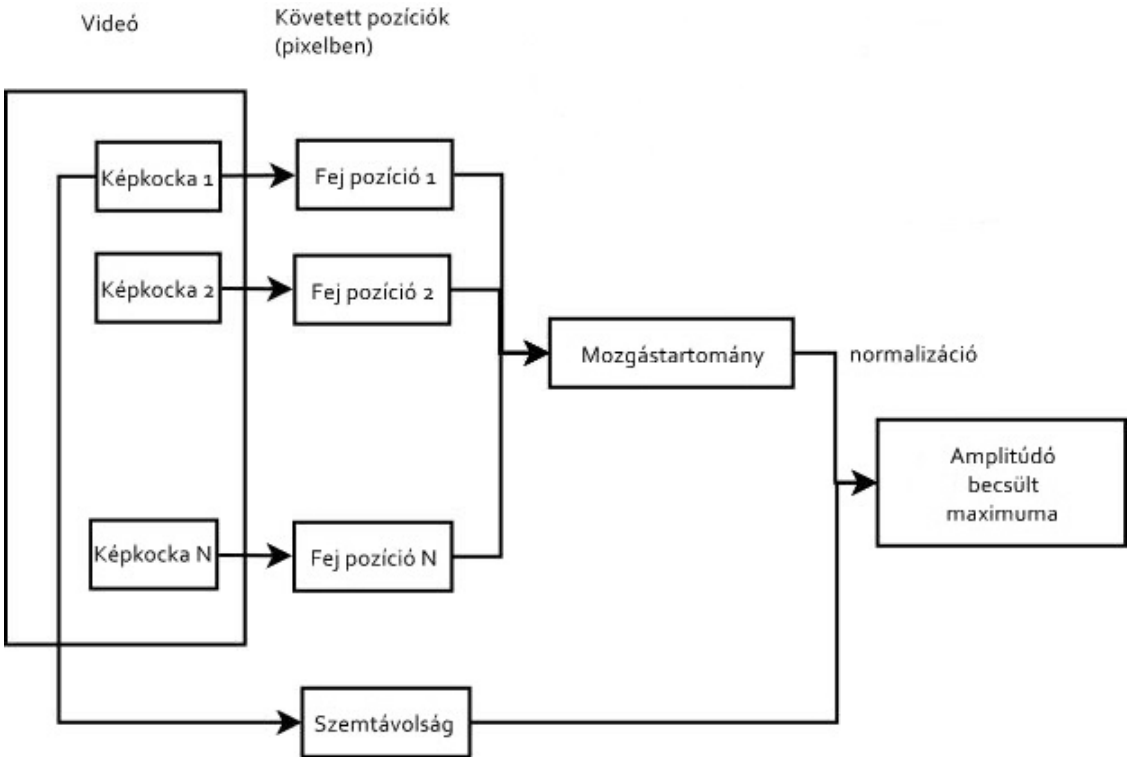
1.2 Mintaadatok, elemzési módszereik és az osztályozás ismertetése

1.2.1 Mintavétel

Az megalkotott osztályozási módszer célja a függőleges fejmozgás amplitúdóinak osztályozása a hanginformációból, a határokat 15, 25 és 50 képpontban határozva meg. Mértékegység tekintetében azért esett pixelre a választásom, mivel egyik fő cél volt, hogy az osztályozás sokféle alany/beszélő esetében lehetővé váljon, amelyhez nagyszámú alany szerepeltetése volt szükséges, ami nem lett volna kivitelezhető a szögek mérését lehetővé tevő, kontrollált körülmények között, ráadásul az adatbázis, illetve a módszer további bővítésének lehetőségét is nagyban behatárolta volna. Ezen túl, az ellenőrzött környezetben történő mintavétel a heterogenitást – és ezzel a módszer határait – szűkítette volna. A szabályalkotó minták

megszerzéséhez – az előző állításokat figyelembe véve – YouTube-videókat használtam bemenetként, amelyekben heterogén (különböző korú, nemű) alanyok spontán beszéltek a kamera előtt, akiknek fejmozgása még természetesebb volt amiatt, hogy mérőkorongok sem szerepeltek fejükön.

A videók minden mondatából 1-1 mintát nyertem ki, azzal az elvárással, hogy a beszélő által a kiválasztott mondat elmondása közben végzett bármilyen testmozgás lényegesen kisebb legyen (pixelben kifejezve), mint a fejmozgás az adott mondat alatt. Az osztályozási módszer kidolgozásához 450 minta került felhasználásra hang- és vizuális információkkal. Minden minta egy adott videó egy mondatából kivont adathalmazból állt, amely megfelelt a korábban leírt feltételeknek. Az 1. ábra a videóinformáció kinyerési folyamatot ismerteti ezen minták esetében.

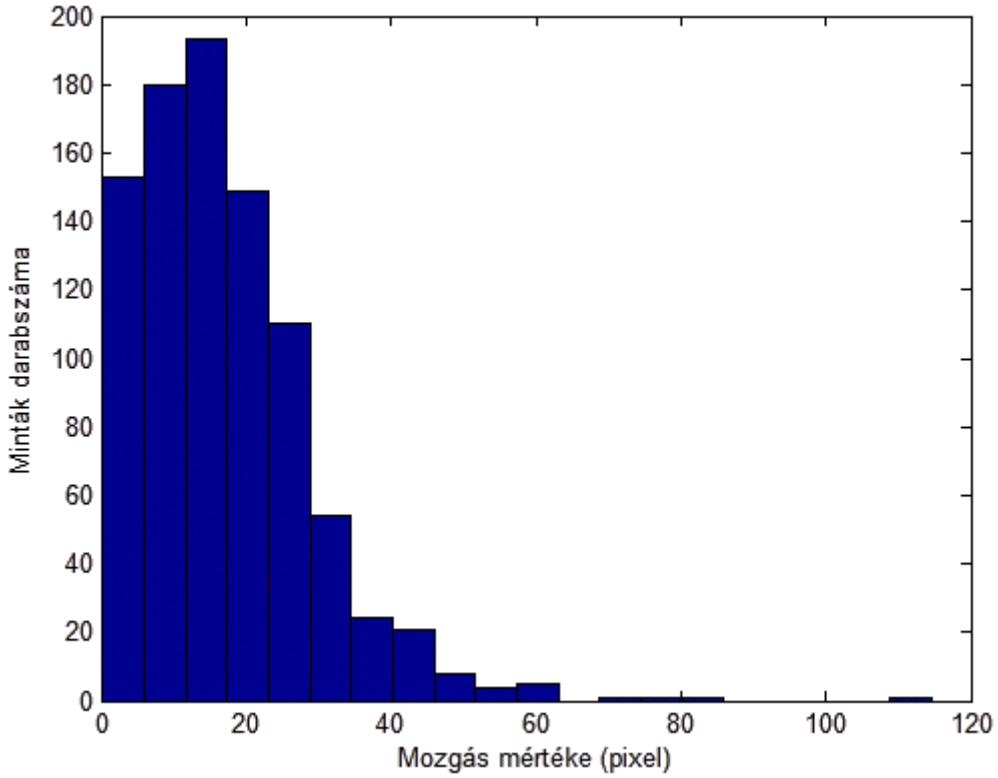


1. ábra. Információkinyerés az osztályozás során

A fentieken túl egy további, körülbelül 450 mintás, ugyanilyen jellemzőkkel bíró mintahalmaz (heterogén alanyok, spontán beszéd, minden minta egy mondat stb.) készült el tesztelési céllal. A két halmaz együttesére a mondatok átlagos időtartama 6,32 másodperc volt. A teljes mintaadatbázisban egy alany legfeljebb 5 mintában szerepelhetett, a beszélők esetleges sajátos jellemzőiből adódó hibák elkerülését biztosítandó.

1.2.2 Mintaelemzés

A függőleges fejmozgás maximumát egy szemmozgás-követő programmal adtam meg, amelynek eredményeit képkockánkként manuálisan korrigáltam. Ezt követően a mozgásokat a szemek közötti 50 pixeles távolságra skáláztam, így kapva meg a fejmozgások mértékének a 2. ábrán látható eloszlását.



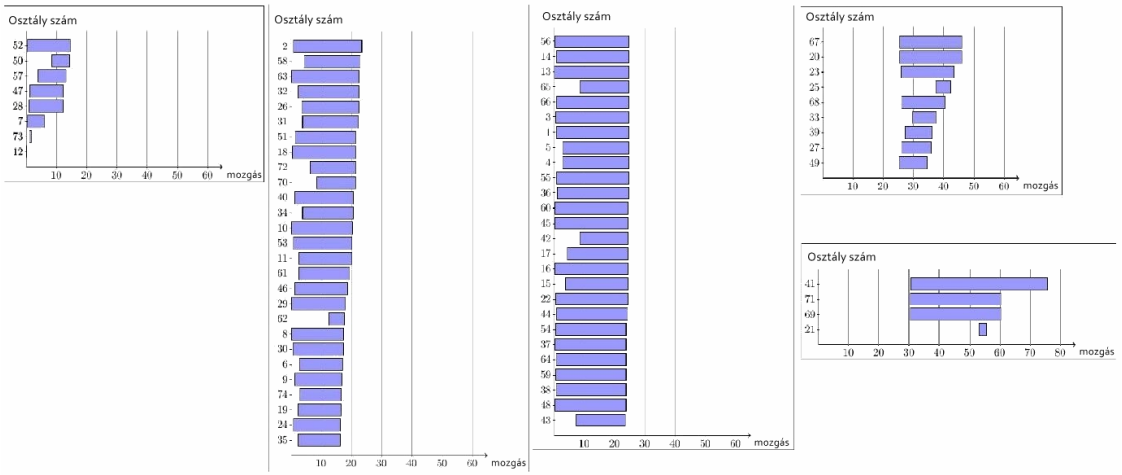
2. ábra. Mozgásintervallumok az osztályozási eljárás mintáinak halmazában

Az osztályozás hangbemenetei részben közvetlenül kinyert értékek voltak (például az intenzitás gyors Fourier (FFT)-transzformációjának varianciája (P_1), autokorrelációs pitch átlaga (P_2), amplitúdó különbségek átlaga (P_3), intenzitás különbségek átlaga (P_4), keresztkorrelációs pitch átlaga (P_6), az amplitúdó varianciája (P_{17}), intenzitás varianciája (P_{19}). Másrészt más hanginformációs értékeket empirikus alapon vezettem le az előbbiekből. Így előállt alap- és a származtatott hanginformációs értékek száma összesen 29 volt. Ezeket együttesen használtam fel az osztálymeghatározó produkciós szabályokban. A kinyert tulajdonságok a Függelék 1. táblázatában láthatók. A konstansok célja az értékeknek a többivel azonos intervallumba transzformálása volt.

1.2.3 Az osztályozási módszer ismertetése

A doktori kutatásaim során kifejlesztett osztályozási módszer hiperterek segítségével osztja fel a jellemzőteret, ami a szakértői rendszerek jól ismert megközelítése. A figyelembe vett pontok halmazát a jellemzőtérben ekkor produkciós szabályok meghatározásával határozhatjuk meg. A szabályok ekkor valójában a konjunkciós operátorok által kifejezett félterek metszéspontját adják meg.

A fent említett 29 hanginformációs tulajdonsággal a teljes mintaadatbázisból osztályalkotási szabályokat készítettem, ahol az egy osztályhoz tartozó tulajdonságok értékei (jellemzően osztályonként 4-5 a rendelkezésre álló 29 tulajdonságból) egy minimum-maximum tartományban vannak. Az egy adott osztályba tartozó minták közös jellemzője az is, hogy maximális mozgásamplitúdójuk egy adott tartományon belül van. Összesen 74 osztályt határoztam meg ezzel a módszerrel (3. ábra).



3. ábra. Az osztályozás során létrejött mozgásmérték osztályok és mozgásintervallumaik

Legyen P_j egy tetszőleges tulajdonság, $j \in N$. Használjunk összehasonlító operátorokat állítások definiálására a megfelelő hipersíkot érintően. Például $P_j < 20$ azt jelenti, hogy az állítás igaz, ha $P_j < 20$. Definiáljunk egy produkciós szabályt, mint ezen állítások logikai formuláját, $f: P \rightarrow \{igaz, hamis\}$. A produkciós szabály a becsült osztályokra is utal. Ezért a szabályaink halmazát úgy formalizálhatjuk, mint (f_i, c_i) párok R halmazát, ahol a f_i egy logikai formula az audiójellemzők tartományán, és c_i a megfelelő osztály. Az osztályozási folyamat során egy minta több osztályba is tartozhat.

A következő lépésben a 0-15 pixel, 0-25 pixel és 25 pixelnél nagyobb mozgási amplitúdók eseteire 3 fő mozgáscsoportot határoztam meg, amelyek a korábban leírt 74 osztályt

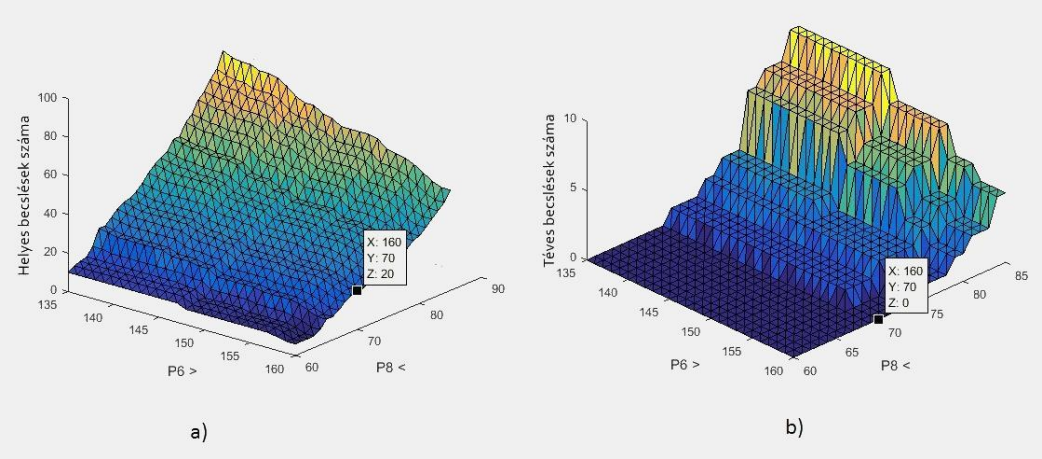
tartalmazzák. A negyedik – 50 pixelnél nagyobb – mozgásamplitúdójú csoport ebben a szakaszban csak néhány mintát tartalmazott, ezért nem vettem be a vizsgálatba.

Ha egy hangminta tulajdonságértékei megfelelnek a 74 kisebb (al)osztály bármelyikének szabályainak, akkor a mintát az osztályba tartozónak tekintettem. Egy alosztályba tartozás ténye becslést ad a minta maximális függőleges fejmozgására, mivel minden alosztály a három nagy mozgáscsoport egyikéhez tartozik. Ezek összegével a módszer jelenleg a minták 76 %-ánál ad becslést a maximális függőleges mozgásra.

Például, ha egy mintából kinyerhető hanginformációs értékek $P_1 - P_{32}$ (amint az a függelék 1. táblázatában látható), akkor amennyiben $P_6 \geq 160$, $P_8 < 70$, és $P_4 > 200$, úgy a minta a 74-ből a 17. alosztályba, az pedig a 25 pixelnél kisebb méretű mozgások főcsoportjába tartozik.

A produkciós szabályok a 74 osztályból 13-ban tartalmaznak "vagy" és "és" operátorokat, a többi osztályban pedig csak az "és" operátort.

Természetesen minden osztályozási módszer érzékenyen reagál a kieső értékekre. Jelen esetben egy mintát akkor tekintek outliernek, ha az bár megfelel egy adott alosztály szabályának, azonban a mozgásintervallum tekintetében kiugró (azaz a mozgás tartománya nem tartozik annak a főcsoportnak a mozgásmértékéhez, amelyhez az alosztály tartozik). A szabályokban szereplő, az adathalmaz megfigyelésével meghatározott konstansok értékeit gépi úton korrigáltam a szabálybázist alkotó mintákra azzal a kívánalommal, hogy értékükkel a lehető legkisebb számú osztályozási hiba (outlier) legyen a halmazon [S3]. A kiigazítás példája a 4. ábrán látható.



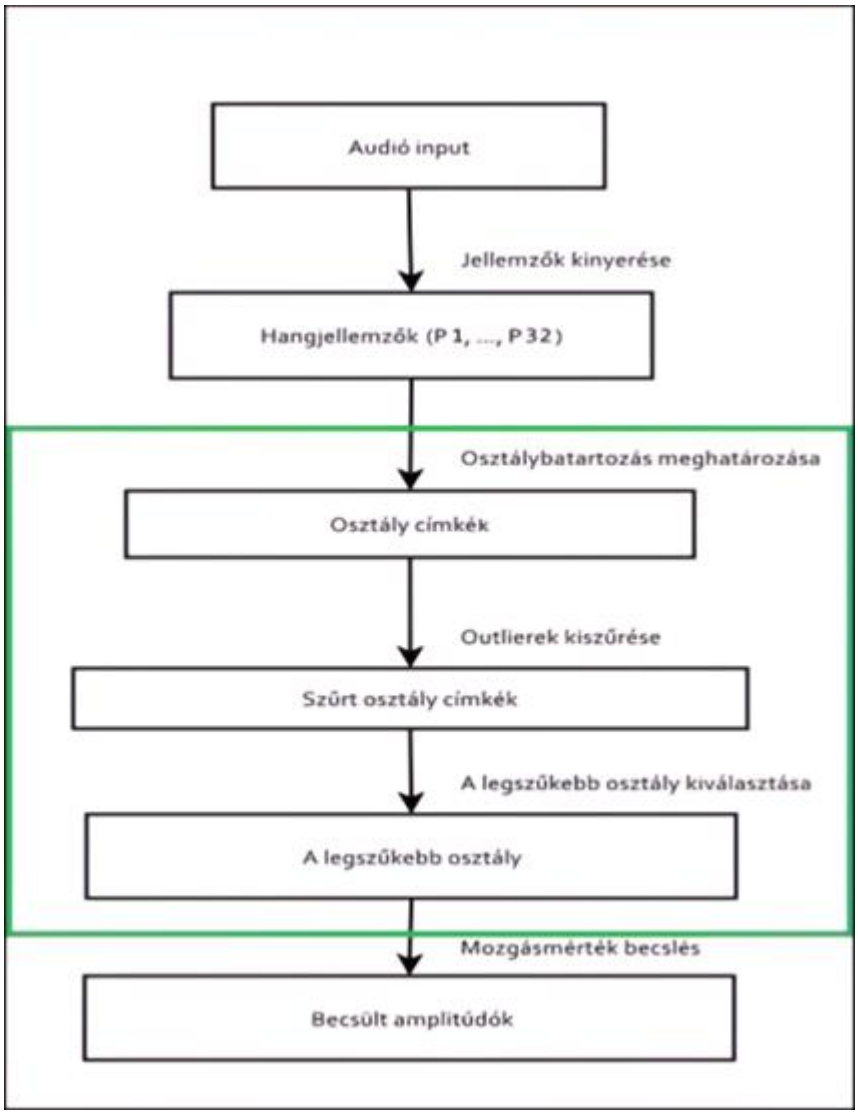
4. ábra. Példa egy osztályszabály konstansainak beállítására

Míg a 4a. ábra egy példán keresztül a helyesen becsült minták számát mutatja az osztálytulajdonságok tényleges értékével, addig a 4b. ábra a tévesen besorolt minták számát mutatja ugyanezen tulajdonságértékekkel. Látható, hogy az értékek kiválasztása során a $P_6 \geq 160$, és $P_8 < 70$ ésszerű, mivel a helyesen becsült minták száma az esetben eléri a 20-as értéket, míg a tévesen besorolt minták száma továbbra is a minimumon marad. A kiigazítást ugyanezzel a módszerrel végeztem el a többi 73 osztály esetében is.

A három főcsoport alatti kisebb, egymással átfedésben álló 74 osztály az egyes jellemzőkben – figyelembe véve, hogy egy minta egynél több osztályba is tartozhat – hasonlóak a biklaszterezés [8] vagy a kapcsolt kétirányú klaszterezés [18] módszereiben használt entitásokhoz. Ezen módszereknek és az általam megvalósított módszernek is közös jellemzője, hogy hasonló jellemzőkkel rendelkező minták részhalmazait vagy csoportjait találják meg, ahol az adott csoport a rendelkezésre álló jellemzők egy részhalmazának intervallumával írható le. Esetemben a hasonló jellemző minden ilyen csoportban a mozgás amplitúdója.

A fentiek szerint kialakított csoportok mozgásértékein egy skálázást hajtottam végre. A 0-15 pixeles főcsoportba tartozó osztályba tartozó mintákat 7 pixelre, míg a csak a 0-25 pixeles főcsoportba tartozó osztályba tartozó mintákat 13 pixelre skáláztam (az osztályalkotó minták mozgásának mediánja). A 25 pixelnél nagyobb mozgáscsoportba tartozó minták 36 pixelre lettek skálázva. Ha egy minta egy olyan osztályba tartozik, amely a 0-15 pixeles mozgásméretű főcsoportba tartozik, de emellett egy olyan osztályba is amely a 0-25 pixeles főcsoportba tartozik, akkor a szűkebb csoportot tekintettem relevánsnak (mivel értelemszerűen ez adja a pontosabb mozgásmérték becslést).

Az 5. ábra áttekintő formában bemutatja az osztályozási folyamat fent ismertetett lépéseit.



5. ábra. Az osztályozás folyamata

1.3 Eredmények és a kialakított rendszer vizsgálata

1.3.1 A pontosság és az újdonság jellemzése

Néhány korai fejmozgás-szintetizáló módszer 1-2 alany fejmozgás mintáit klaszterezte [11], és minden egyes klaszterhez Rejtett Markov Modellt tanítottak [21], [20]. Korábbi kutatások azonban kimutatták, hogy a mozgásgenerálás első lépéseként a képkockák szerinti klaszterezés nem a legjobb megközelítés [35]. Az általam megvalósított osztályozási megoldás – a fentiek szerint – nem ilyen módon alkalmazza az osztályozást. Fontos fejlesztési eredményem emellett az is, hogy jellemzői okán bármely fejmozgás-generálási eljárás kiegészíthető az osztályozási eljárásommal (végső lépésként), amennyiben a generált mozgásamplitúdók skálázhatók egy, az osztályozási módszerrel számított konstans értékkel.

Voltak más kísérletek is (bár nem ezzel a céllal) a függőleges fejmozgás amplitúdójának implicit becslésére. Ezek pontossága a generált mozgás amplitúdójának és a referenciának az arányából számítható [63], [19]. Közös jellemzőjük azonban, hogy a generált fejmozgásokat néhány alany segítségével konstruálták, és hogy a szintetizált mozgás amplitúdója és a referencia közötti szorzó egy mondatos intervallumban a bemutatott példákban is 2-10 között változott. Ez az arány akár a 60-at is elérheti [24], és Matthews és szerzőtársai legújabb eredménye [35] példájában 7 körül adódott. Fenti szerzők jövőbeli célja, hogy módszerük általánosítási lehetőségeit keressék, és a korpuszukon kívüli beszélőket is becslést adjanak. Egyes módszereknek a becslés kezdetétől [19] vagy a fej korábbi pozíciójából [21] származó referencia vizuális információra is szükségük van. Zhou és társai [66] egy olyan mozgásgenerálási módszert építettek fel, amely különbözik a fent említettektől, mivel csak egy vizuális képkockát használ, azonban csak az ajak (és környéke) mozgásgenerálására koncentrál, és egyéb korlátozásokkal rendelkezik. Ezzel szemben az általam javasolt osztályozási módszer különböző alanyokra épül, és nem használ vizuális információt a fejmozgás becsléséhez. A módszer általánosan alkalmazható voltához hozzájárul a fenti jellemző, miszerint a rendszer különböző alanyokra épül fel, és tesztminták által bizonyítottan a korpuszon kívüli alanyok esetében is működik azokban az esetekben, amelyekben a hangérték beleesik valamely osztályba. Kivételt képeznek ez alól a magas érzelmi töltésű esetek – ez a módszer továbbfejlesztésének egy potenciális iránya lesz. A megoldás jelen változatának legnagyobb becslési hibája – helyes osztályba sorolás esetén – a más, ismertetett módszerekkel összemérhető 8-10-es szorzót érheti el legrosszabb esetben, aláhúzandó viszont, hogy mindezt – a többi megoldással ellentétben – vizuális információ felhasználása nélkül elérve. Ezen említett maximális hibaszorzó abban az esetben áll fenn, ha a mondatban a legnagyobb amplitúdó minimális (1-2 pixel), miközben a minta csak a 0-25 pixeles mozgásintervallum-csoport egyik osztályába sorolható, miközben a 0-15 pixeles mozgásintervallum-csoport egyik osztályába sem. Az amplitúdók ez esetben ugyanis úgy kerülnek skálázásra, hogy a maximális amplitúdó érték 13-ra adódjon (a 0-25 pixeles csoport miatt). Ilyen esetek a minták kevesebb, mint 12%-ában fordulnak elő. Továbbá az esetarány tovább csökkenthető, ha több osztályt határozunk meg a 0-15 pixeles (vagy még szűkebb) tartományban – csökkentve annak az esélyét, hogy egy mindössze néhány pixeles mozgást mutató minta csak a 0-25 pixeles intervallumcsoportba kerüljön.

Egy további eredmény a területen [29] egy olyan kutatócsoporthoz köthető, akik egyrészt felhasználnak szájmozgást (mint az előző példában), illetve egyes kiválasztott alanyok

fejmozgását is, és ezekből általánosítanak. Ez több szempontból természetes mozgást eredményez, azonban például a tanítóminták közötti nagy fejmozgások (és feltehetően egyéb speciális jellemzők) egyes esetekben természetellenesen nagy mozgást eredményeznek. Megjegyzendő, hogy ez javítható lenne az általam megvalósított osztályozási megoldás korrekciós fokozatként történő alkalmazásával azon módszer végső fázisaként.

A fentieket összegezve, módszeremben a generált mozgás és a referencia amplitúdójának aránya összehasonlítható a legújabb más kutatásokkal, miközben – azokkal ellentétben – nem használ vizuális információt (amely nagyban korlátozná felhasználhatóságát), az alanyok tágabb körén működőképes, továbbá a megoldás nem korlátozódik a jelenlegi állapotára, hanem új osztályok meghatározásával továbbfejleszthető.

A vizuális információ nélküli mozgásgenerálási megoldások közül egyik ismert eljárás sem mutat a fentiekkel összemérhető pontosságot ill. hibaarányt. Ezek esetében a túlzott amplitúdó mellett szintén tipikus kimeneti hiba, hogy adott alanyok esetében nullához közeli amplitúdót generálnak [66], [33]. A mozgásgenerálás mindkét típusú hibája csökkenthető az általam javasolt osztályozási módszerrel akár ezen más eljárásokhoz csatoltan is, az adott megoldás kiegészítő fázisaként végzett automatikus amplitúdó korrekciós moduljaként alkalmazva. Emellett a vizuális információt használó megoldások kimenetének javítására is ugyanígy felhasználható.

A fentieken túlmenően vizsgáltam módszerem egyéb továbbfejlesztési lehetőségeit is. Ilyen például, hogy amennyiben egy minta a 74-ből több osztály szabályainak is megfelel, akkor eredményként megadhatjuk ezen osztályok legrövidebb mozgásintervallumának közepét. Ezzel a módosítással a 0-25 pixeles mozgásintervallum-csoportba tartozó osztályok mintáinak 14%-ánál, a 25 pixelnél nagyobb mozgásintervallum-csoportba tartozó osztályok mintáinak 42%-ánál a lehetséges mozgásintervallum tovább zsugorítható, mindkét esetben több, mint 12 pixellel. Bár ez a változat egy plusz szükséges lépést ad az amplitúdóskálázáshoz, nagyobb pontosságot eredményez.

1.3.2 Az egész osztályra kiterjedő erős korreláció és annak következménye

A legalább 10 mintával rendelkező 42 osztály esetében az osztályt alkotó mintákra felírható egy az osztályra jellemző paraméterarány (pl. $P_{27}/P_4/P_{19}$), ahol az arány értéke és a mozgás közötti lineáris korreláció az osztály elemeire jellemzően 0,9 feletti. Ennek a korrelációs együttható értéknek az átlaga a 42 osztályra vonatkozóan 0,8758, a maximuma pedig 0,9979 volt vizsgálataimban.

Ennek gyakorlati jelentőségét egy példa kapcsán ismertetem a következőkben. Tartozzon az ismeretlen mértékű mozgással rendelkező hangminta abba az osztályba, ahol a 0,9-es korreláció a mozgás és a $P_{27}/P_4/P_{19}$, az osztályra jellemző arány adta érték között. Ebben az esetben az arány értéke és a lineáris kapcsolat pontosabb mozgástartomány-becslést ad az adott minta számára, mint amennyit az osztályba tartozás önmagában adna.

A fentiekkel szemben olyan további jellemzőt tapasztaltam vizsgálataim során, hogy amikor ugyanazokat a tulajdonságokat a produkciós szabályrendszer két különálló osztályának egyesítése után vizsgáltam, a korrelációs érték 0,03-0,1 közé esett, és az adatbázis egészét tekintve jellemzően 0,05 volt. Ez a jellemző pedig a jelenlegi osztályozás és az azon belüli egyes osztályok érvényességét támasztja alá.

Ben Youssef és szerzőtársai [5] a CCA (Canonical Correlation Analysis) segítségével 0,2 körüli korrelációt ért el a beszéd és a fejmozgás jellemzői között a kisszámú alany esetében. Egyes kutatók [24] korábban azt állították, hogy általános jelleggel csak 0,07-0,08 körüli korrelációs érték mutatható ki hang és fejmozgás között. Yehia és társai [63] ezzel szemben erős korrelációról számoltak be az intenzitás és a fejmozgás között, azonban Ben Youssef és munkatársaihoz hasonlóan csak kisszámú alany esetében, ráadásul a kapcsolat erőssége mondatról mondatra változott, és a fejmozgás becslése az intenzitásból rosszabb eredményeket mutatott, mint az intenzitás becslése a mozgásból (átlagosan 0,37 korrelációs együttható). Ezeket az eredményeket figyelembe véve, összegzésként megállapítható, hogy a beszéd- és mozgásjellemzők közötti képkockánkénti globális korreláció gyenge.

A korábbi szerzők korrelációs módszere (CCA - kanonikus korrelációs elemzés) eltér az esetemben használt Pearson-féle korrelációs mátrix számításától, ráadásul én a maximális mozgás amplitúdójával összevetésben vizsgáltam az adatokat. A gyenge „globális” (azaz alapvető hangjellemzőkre a teljes adatbázisra igaz) korrelációra vonatkozó megállapítás azonban egybevág az általam kapott eredményekkel, hiszen az én tapasztalataim is azt mutatták, hogy az adatbázis-szintű általános kapcsolat az alapvető beszédjellemzők és a mozgásamplitúdó között gyenge (jellemzően 0,05). Ezzel szemben azonban én származtatott beszédinformációk felhasználásával osztályszintű erős korrelációt találtam, ezzel kimutatva, hogy meghatározhatók olyan származtatott hangjellemzők, amelyekkel adott csoportba tartozó minták esetében ezek, és a minták mozgása között erős lineáris korrelációs kapcsolat van (legalábbis a maximális mozgásmértékre vonatkozóan).

Mindezen felül, ez megerősíti az általam létrehozott/alkalmazott, származtatott hanginformációs jellemzők és az ezeken alapuló osztályok létjogosultságát is. Hozzátéve, hogy további osztályok meghatározása a rendszer jellemzőiből adódóan tovább növelné a mozgásamplitúdó-becslési pontosságot nem pusztán a korrelációs kapcsolat fent ismertetett felhasználásából adódó, de az ettől független, csak az osztályba tartozáson alapuló becslési módszer esetében is. Megjegyzendő itt, hogy ezzel ellentétben az osztályok összevonása nyilvánvalóan csökkentené azt.

1.3.3 Az osztályok száma csökkentésének vizsgálata

Mivel az ismertetett módszer szerint az osztályba sorolások az előre definiált szabályok alapján automatikusan meg tudnak történni, így nem jelent problémát az osztályok száma. Ugyanakkor célszerű volt elemezni ennek lehetőségét, illetve ennek hatását.

1.3.3.1 Első módszer

Az elemzés során minden két osztályt minden lehetséges kombinációban párosítottam, és minden egyes közös páron klaszterezést és validáló eljárást végeztem. A klaszterezési eljárás előtt a párokon dimenziócsökkentést futtattam le.

Az összekapcsolt mátrix-párokon végzett klaszterezést megelőző dimenziócsökkentő technika a Landmark Isomap [44], ennek kimenetén pedig NPE volt (tehát két dimenziócsökkentő módszert alkalmaztam). A két eljárás dimenziószáma három és kettő volt (nagyobb dimenziószám nem volt lehetséges, mivel a vizsgált bemeneti osztályok változóinak száma jellemzően 4-5).

A csoportosítási módszer a két dimenziócsökkentő eljárás után az single linkage a legrövidebb távolság módszerével és az euklideszi távolság volt.

A vizsgálat folyamata: Legyen a ténylegesen vizsgált pár mátrixa *A* és *B*. Ezt a ténylegesen vizsgált két mátrixot összekapcsoljuk. Legyen az összekapcsolt pár mátrixa *C*, és a dimenziócsökkentett *C* mátrix *D*. Minden ilyen *D* mátrix vizsgálatakor explicit módon megadunk egy csoportszámot a single linkage eljárásnak. Ez a csoportszám 1-től 15-ig fut a vizsgálatban.

Az optimális entitásszám az a csoportszám lesz, amelyet a Gap validációs index [58] optimálisnak jelez. Ha az optimum egy, az azt jelenti, hogy a két, vizsgálat okán összekapcsolt osztály egyesíthető. Ezzel szemben, ha az optimum például 7, azt jelzi, hogy az aktuálisan vizsgált kétosztályos adatok 7 részre csoportosítása optimális az index szerint. Ez az optimum a párok 99,7 százalékában 2 és 15 között változott, és csak a fennmaradó 1,3 százalékban volt

1 (és az ilyen esetek csupán 1/3-a tartozott egy adott osztályba). Ezen eredményekből is az a következtetés vonható le, hogy az osztályok összevonása nem ésszerű.

A Gap-validációt elsősorban amiatt választottam, mivel az egy csoportba tartozó adatok összehasonlítását tekintve ez a mutató mutatja meg a legpontosabban, hogy az adatok egy entitáshoz tartoznak [58]. Az index mellett további érv volt, hogy egy másik [16] kimutatta, hogy ez az index a csoportok adatai közötti legrövidebb delta-távolság esetén is a valóshoz legközelebb álló, ill. a legkisebb csoportszámot eredményezte, ami megfelelt a feltételeinek.

1.3.3.2 Második módszer

A fent említett módszer mellett az összes osztálypárt nemcsak a Gap-indexszel, hanem a Davies-Bouldin- [10], valamint a Calinszki-Harabasz-indexszel [6] is vizsgáltam. A vizsgálati módszer – az indexeken kívül – ugyanaz volt, mint korábban. Mind a Davies-Bouldin-index, mind a Calinszki-Harabasz-index minden vizsgált osztálypár esetében a 2-15 optimális csoportszámot jelezte, ami megerősíti a fenti következtetést, hogy az osztályok összevonása nem indokolt.

1.3.3.3 Harmadik módszer

Az olyan klaszterezési/csoportosítási technikák esetében, amelyek a tér sok részhalmazát eredményezik (mint megoldásomban), a páros validálás egy gyakori validálási technika. Ezen túl azonban az eredményeket úgy is megvizsgáltam, hogy a teljes mintaadatbázist egy ismert eljárás bemeneteként használtam. Ebben az esetben mind a single linkage, mind a k-középpontú klaszterelemzési (k-means) eljárás a bemeneti paraméter által megengedett maximális entitásszámot eredményezte, szinte az összes fent említett index és kritérium szerint. Két kivétel közül az egyik volt a sziluett validáció euklideszi távolsággal; azonban még ebben az esetben is 5 volt az optimális csoportszám, és a sziluett azonos számú és méretű pozitív és negatív csúcsokat tartalmazott. Így ez a kritérium nem adott jó eredményt. A másik kivétel a teljes adatbázist két mintától eltekintve egy csoportnak tekintette. Ráadásul ez a megoldásomhoz hasonló hibaarányú mozgásosztályozást sem produkált. Ez a hibaarány – amely módszeremben jelenleg 13 % – talán a legfontosabb mérőszám, amikor a produkciós szabályrendszert bármely más eljárással összehasonlítjuk.

1.3.4 A validálási eljárások eredményeinek összehasonlítása

Az első és második vizsgálati módszerben leírt lehetséges kontrakcióelemzést figyelembe véve fontos megjegyezni, hogy a három index közül kettő nem jelölte optimálisnak az explicit kontrakciókat, míg az egyik index csak az esetek 1,7 százalékában jelölte optimálisnak az

összevonást. Továbbá, ha egy potenciális szűkítés bármelyik módszer vagy index szerint lehetséges vagy optimális, akkor annak nem lehet más előnye, mint a kisebb osztályszám, miközben lehetséges hátrányai is vannak. Ezek közül az első a produkciós szabályok hosszának növekedése. A második a megoldás lehetséges továbbfejlesztésének megvalósulása esetén merülhet fel. Ebben az esetben ugyanis – a főcsoportok használata nélkül – a mozgástartományt csak a 74 alosztályba tartozás határozná meg, összevonással pedig ez a tartományszélesség növekedne, így a becslési pontosság ebben az esetben romlana.

A harmadik módszer, amely a teljes adatbázist adta bemenetül a jól ismert csoportosítási technikákba, nem tudott az enyémhez hasonló eredményeket produkálni.

1.3.5 A változók számának vizsgálata

Az osztályok szabályrendszereit alkotó változók számának (29 alap- és származtatott adat együttesen) lehetséges csökkentése érdekében először két lineáris módszert, főkomponens-analízist [25] és faktoranalízist [45] vizsgáltam.

Mivel a faktoranalízis nem vezet be új, független változókat, hanem az eredetiek értelmezését teszi lehetővé azok közös faktorai révén, a produkciós szabályok egyszerűsítésének egyik lehetősége, hogy változóikat a közös faktorokkal helyettesítjük. Emellett, ha a szűkítésre jelölt szabálypárok változóit faktorokkal írjuk fel, lehetséges osztály-összevonások is könnyebben láthatóvá válhatnak.

Ebből a célból faktoranalízist alkalmaztam azokra az osztályokra, amelyek legalább 7 mintát tartalmaznak. Az osztályonként jellemzően 4-5 változóból a specifikus variancia minden osztályban legfeljebb egy változó esetében volt 0,1 alatt, tehát a változók varianciájának nagy százalékát minden osztályban csak egy esetben lehet ezzel az eljárással lefedni. A specifikus varianciák osztályonkénti átlagainak minimuma 0,29 volt, míg a teljes specifikus variancia átlaga 0,5367-ként adódott. Ezek okán a változószám faktoranalízissel történő csökkentése nem ésszerű.

Főkomponens-analízissel (PCA) az osztályok változóit osztályonként 1-2 főkomponenssel lehet lefedni. Hátrányként merül fel azonban, hogy vizsgálataim szerint a rekonstruált értékek 25-100%-os eltérést mutatnának az eredetitől (ami túlzott hibát adna a rendszerhez), és az eljárás főkomponensként értelmezhetetlen változókat is bevezetne. Így a PCA alkalmazása erre a rendszerre nem ésszerű.

Néhány nemlineáris módszer esetében – GDA (Generalized Discriminant Analysis) [4], NPE (Neighborhood Preserving Embedding) módszer [32] és LLC (Locally Linear Coordination)

[40] – azt vizsgáltam a szokásos paramétereikkel, hogy a kimenetükön k-means CA vagy linkage eljárások alkalmazásával megvalósítható-e a hanginformációkból történő a mozgásmérték osztályozás. A dimenziók száma mindhárom eljárás esetében 2, 4, 6 és 8 volt. A GDA esetében a kiegészítő paraméter a Gauss-kernel volt, míg az NPE esetében 12 szomszéd. Az LLC módszer – Locally Linear Coordination – alkalmazásakor a paraméterek 12 analizátor és 200 iteráció voltak. Értékelési, validációs kritériumokként minden esetben a Gap [58], Davies-Bouldin [10] és Calinszki-Harabasz [6] eljárásokat alkalmaztam. A módszerekkel nem sikerült külön mozgásmérési osztályokat létrehozni, kivéve a GDA dimenziócsökkentést, amelyet linkage vagy k-means csoportosítás követett sziluett vagy Davies-Bouldin validációval. Ezekben az esetekben, különböző GDA dimenziószámok mellett, több mint 100 mintából álló csoport alakult ki, jellemzően 20 alatti mozgásmértékkel. Ez azonban – az ismertetett osztályozási megoldással ellentétben – nem tudott ezen az egyen kívül más mozgásintervallum osztályokat létrehozni, és nem tudott annyi mintát csoportosítani (~30%).

1.3.6 Tézis

A fejezetben ismertetett eredmények alapján az alábbi új tudományos eredményeket fogalmaztam meg.

1. Tézis: Kidolgoztam egy kizárólag hanginformációkat felhasználó, szabályalapú osztályozási módszert, amely képes a fejmozgás mértékét kellő pontossággal becsülni. A módszer a betanításban nem szereplő alanyok mozgásának tartományát is képes becsülni. Bebizonyítottam, hogy képezhetők olyan hanginformációs paraméterek, amelyek mind a mozgásmérték osztályozását, mind pedig a hang és a fejmozgás közötti korrelációt tekintve jobb eredményt adnak, mint a hagyományosan alkalmazott hanginformációs jellemzők. Igazoltam, hogy az általam megalkotott osztályozási eljárás jelen struktúrájában megfelelő megoldást ad. A módszer az általam vizsgált esetekben jobb megoldást ad, mint a dimenziócsökkentéssel támogatott változat, illetve a gyakorlatban hagyományosan alkalmazott eljárások.

Kapcsolódó saját publikációk: [S1, S2, S3]

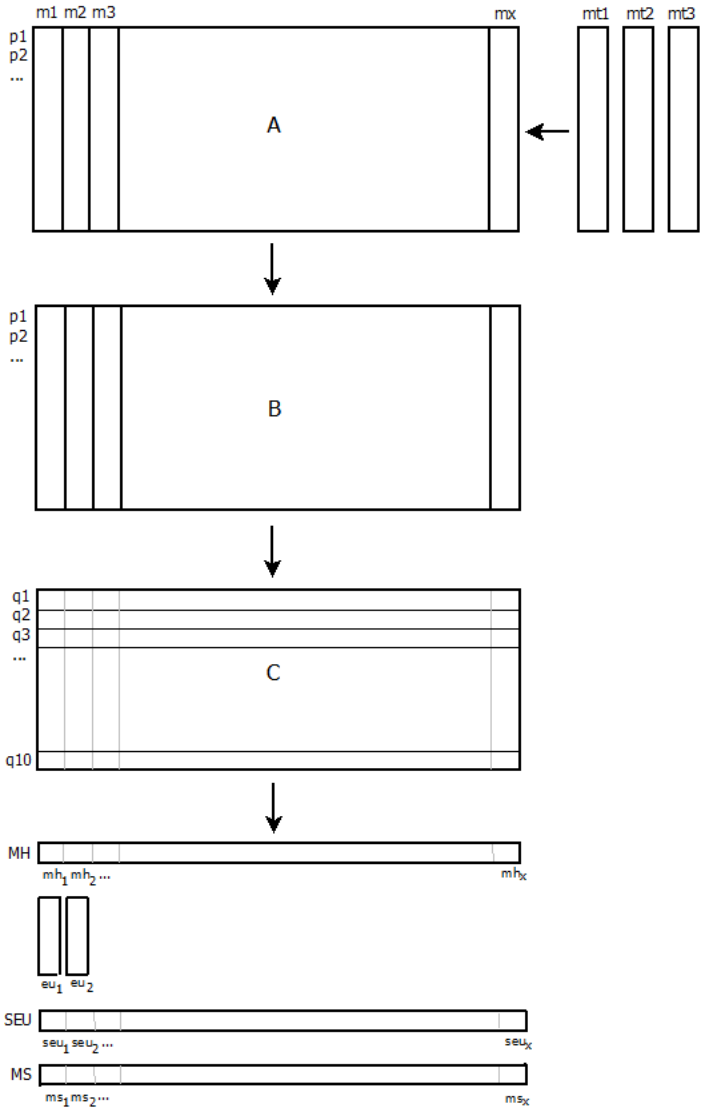
1.4 Outlierek észlelése

Egy minta akkor tekinthető egy osztályon belül kiesőnek (outliernek), ha hanginformációs értékei megfelelnek az adott osztály szabályának, azonban mozgást tekintve nem tartozik az osztály mozgási intervallumába. Az ilyen minták esetében a rendszer téves mozgásamplitúdó-

becslést adna, ezért automatikus kizárásuk az osztályszabályban nem szereplő jellemzőik alapján indokolt. E cél elérése és a fenti eredmények további javítása érdekében a következő automatikus outlier-detektálási mechanizmust alkottam meg.

1.4.1 Outlierek észlelésének szabálybázisának meghatározása

Az osztályozási megoldás jelenleg 905 mintából 217 esetben nem tud becslést adni. A többi, jelenleg kezelt 688 mintát illetően az esetek 87 százalékában helyesen osztályoz, 13 százalékban pedig tévesen. Outliernek ezen minták tekinthetők. A fent említett 13 százalékos téves besorolási arány az outlierek észlelése nélkül értendő, amit a következő outlier detektálási eljárással csökkentettem. A javasolt detektálási eljárás a 6. ábrán látható.



6. ábra. A kiugró értékek felismerésének eljárása

Az első lépésben veszünk egy *A* osztálymátrixot, amelynek elemei $m_1, m_2, \ldots$ Ezt kiegészítjük a $mt_1, mt_2, \ldots$ outlier detektálás-tanulási mintákkal, amelyek szintén megfelelnek az osztály

szabályának. (Ezen tanítómintákból néhány outlier, néhány nem.) Így rendre egy B mátrixot konstruálunk minden *A* mátrixból, azt egy *mt* mintával kiegészítve. Ezekben a mátrixokban a p_1, p_2,... hanginformációs értékek (tulajdonságok) közül minden lehetséges kombinációban 10-et választunk ki, C mátrixokat megalkotva (a ténylegesen kiválasztott 10 tulajdonság elnevezése: $q1, q2, \dots, q10$).

Az 6. ábrán, amely a kiugró értékek felismerésének eljárását ismerteti, mh_1 a Mahalanobis-távolság C mátrixban a $q1{:}\, q10$ hanginformációs tulajdonságvektorra *m1* oszlop esetében, ahol m1 az első minta. Hasonlóképpen, mh_2 ugyanez a második minta esetében stb. Eu_1 az euklideszi távolság *C*-ben $q1 \dots q10$ első oszlop értékei (1. minta) és a többi minta azonos *q1:q10* indexű értékei között. Eu_2 ugyanez a mátrix második oszlopának értékei esetében stb. Seu_i pedig $std(eu_i)$ és $ms_i = mh_i/seu_i$.

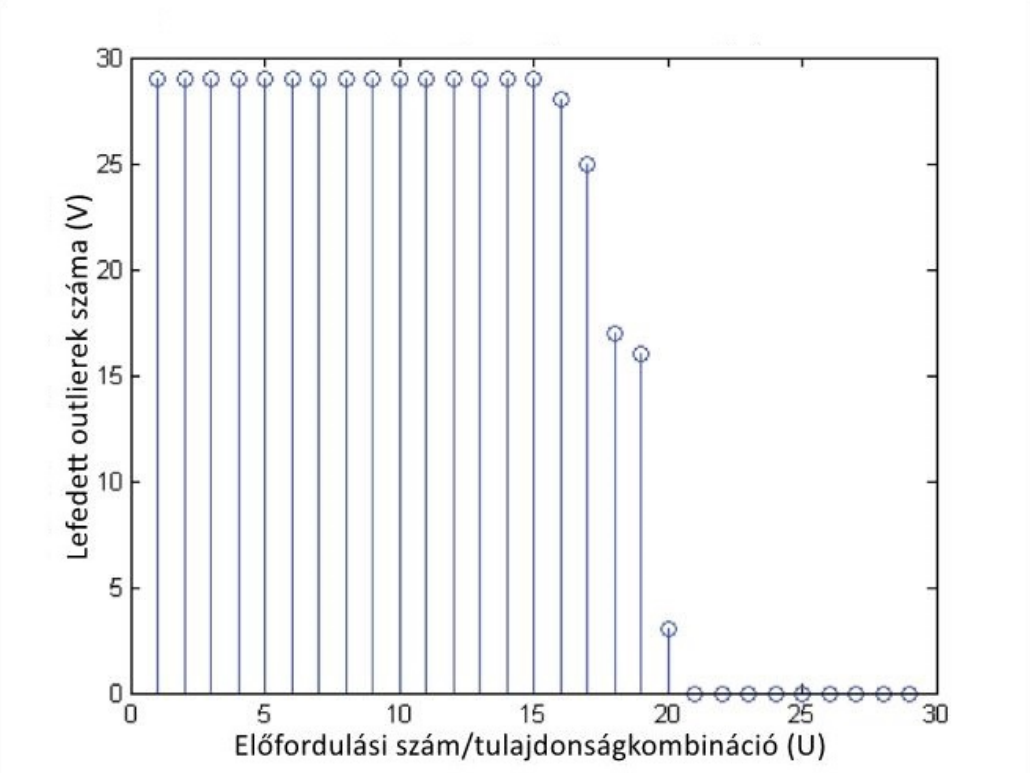
A C mátrixban az ms_1, ms_2 elemek, amelyeket az eredeti osztály mátrixból kaptunk, és ms_x aktuális mintából kapott elem együttesen MS vektort eredményezik. Amennyiben ms_x elem értéke MS minimuma vagy a maximuma, akkor az adott $q1, \dots, q10$ tulajdonságazonosítókat megtartjuk. Ahhoz, hogy egy adott $q1, \dots, q10$ tulajdonságazonosító kombináció (például $p_1, p_3, p_4, p_8, p_{10}, p_{13}, p_{19}, p_{20}, p_{22}, p_{23}$) érvényesnek legyen tekinthető, két feltételnek kell teljesülnie. Az első, hogy több osztályra legalább 4 outlierre legyen igaz, hogy az ms_x elem (amelyet az aktuális elemhez számolunk $q1, \dots, q10$ kombinációval), a minimum vagy a maximum értéket képviseli a saját osztálya által előállított MS vektorban. A második feltétel az, hogy ugyanaz az ms_x ne legyen egyidejűleg minimum vagy a maximum egyik ugyanazon $q1, \dots, q10$ tulajdonságvektorral rendelkező osztályban sem, ahol az adott minta nem outlier. Ha a két fenti feltétel egyidejűleg teljesül, és így egy $q1, \dots, q10$ kombináció érvényes, akkor minden egyes MS vektorra, amelyet ebből egy jövőbeli vizsgálati mintára és az osztály alapmintáira számítunk, a vizsgált minta outliernek tekinthető, ha az ms_x elem minimális vagy maximális MS-ben.

Megvizsgáltam, hogy a módszer hány kiugró értéket tud lefedni a legalább 10 elemű osztályokban, felhasználva azt a körülbelül 450 tesztmintát, amelyet nem használtam fel az osztályszabályok kialakításánál. A 25 pixelnél kisebb mozgásamplitúdójú elemek osztályaiban 29, az ennél nagyobb mozgásamplitúdójú elemek osztályaiban pedig 23 volt a kiugró értékek száma (tehát 29 és 23 volt a maximálisan detektálható outlier).

Azt is megvizsgáltam, hogyan változnak az eredmények, ha nem 10, hanem 5, 7 és 12 hangtulajdonságot tartalmazó kombinációkat választok, valamint azt az esetet is, amikor 10

hanginformációs tulajdonság esetén az MS vektor elemei a következőképpen kerülnek kiszámításra: $ms_i = (mh_i - seu_i)/mh_i$ nem pedig $ms_i = mh_i/seu_i$ formában.

Az 7. ábra a fenti 25 pixeles mozgásmértékű osztályok esetét mutatja outlierek tekintetében. A függőleges tengelyen látható az outlierek száma, amelyeket a kombinációk összessége lefed (V). A vízszintes tengelyen azon követelmény számszerű értéke látható (U), hogy minden tulajdonságkombinációnak legalább hány outlierben kell előfordulnia. Az ábrán látható esetben a kombinációkat 5 hanginformációs tulajdonságból választottam ki ($q1, \dots, q5$), $ms_i = mh_i/seu_i$ MS vektorelemeket használva. Látható, hogy még azzal az – eredetinél szigorúbb – követelménnyel is, hogy minden tulajdonságkombinációnak legalább 15 outlierben kell előfordulnia (miközben nem fordulhat elő egyetlen nem kiugróban sem), a módszer mind a 29 outlier kimutatására képes volt.



7. ábra. Az outlierek észlelési paramétereinek beállítása

A 25 pixel alatti mozgásméretű osztályok esetében a legjobb eredményt szintén $ms_i = mh_i/seu_i$ formában értem el, de 10 elemkombinációval. Ebben az esetben azonban a maximális számú outliert (itt 23) úgy lehet lefedni, hogy a kombinációk elvárásként 5 outliernél fordulnak elő.

1.5 Az eredmények összegzése

A fentiekben a függőleges fejmozgás amplitúdójának hanginformációkból történő osztályozására szolgáló módszert mutattam be. A módszer előnye, hogy – más megközelítésekkel ellentétben – nem igényel előzetes mozgásmintákat az alanytól, és a létrehozás jellegéből adódóan alanyok tág körére is működik.

Az osztályozási megoldás jelenleg a vizsgált minták nagy részében (76 százalékában) adott megbízható becslést a maximális mozgás amplitúdójára pusztán a hang alapján. Ez az arány – az előtanításon alapuló módszerekkel ellentétben – tovább növelhető a tér nem lefedett részeit tartalmazó új osztályok definiálásával, illetve a becslés pontossága növelhető kisebb mozgásintervallum-osztályok megalkotásával. Azokra az esetekre, ahol az osztályozási módszer jelenleg képes becslést adni, a megoldás 87 százalékban jó becslést ad, mindennemű outlier detektálás nélkül. A módszer javításának céljával kifejlesztésre és bemutatásra került általam egy olyan outlier detektálási eljárás is, amely a fent ismertetettek szerint ezt tovább javítja. (Segítségével a helyes becslési arány akár 100 százalékra is növelhető vizsgálataim szerint.)

Az értekezésben ismertettem az általam kidolgozott eljárást, annak használatához szükséges hanginformáció kinyerését, az osztályok létrehozásának módszerét és a csoportok tulajdonságait, felhívva a figyelmet az osztályokon belül elérhető magas hang-mozgás korrelációs értékekre is, amelyek további lehetőségeket kínálnak. A fentieken túl a módszer más megoldásokkal való összehasonlítását, hatékonyságát és az eljárás fejlesztéseit is tárgyaltam a dolgozat előbbi fejezetiben.

Összességében megállapítható, hogy a javasolt osztályozási eljárás az esetek jelentős részében javíthatja a fejmozgás-generáló módszerek kimeneteinek természetességét, azáltal, hogy lehetőséget ad az osztályozás alapján, végső fázisban történő amplitúdó korrekcióra, mindezt olyan formában megvalósítva, amely nem igényel vizuális információt.

1.6 Tézis

2. Tézis: Megalkottam a mozgásmérték osztályozási eljárás hatékonyságának növelése céljából egy a kieső értékeket hatékonyan kiszűrni képes, Mahalanobis és euklideszi távolság együttes használatával operáló outlier-detektálási eljárást.

Kapcsolódó saját publikáció: [S3]

2. Hisztogramszűrésen alapuló súlyozott medián zajszűrési eljárás

2.1Bevezetés

A Leggyakoribb Érték, vagy Most Frequent Value (MFV) módszer [46] egy nagy statisztikai hatásfokú becslési eljárás, mely robusztussága miatt jól alkalmazható zajos mérési adatrendszerek feldolgozásában. Az MFV módszert eredetileg a földtudományban fejlesztették ki, mely a geofizikai [12] [53] [48], hidrogeológiai [54] és számos más földtudományi terület [64] mellett előnyösen alkalmazható pl. a kozmológiában [64], űrgeodéziában [60], stb. Ilyen, különösen szélsőségesen zajos adatrendszerek a térinformatika területén is jelen vannak, beleértve a digitális domborzatmodellezést és a műholdas adatátvitelt. Hasonlóan elterjedt, de kevésbé robusztus és kifinomult technika a mediánszűrés [47] [26].

A potenciális alkalmazási területek közül a digitális domborzati adatok zajcsökkentésére dolgoztam ki új, hisztogram-alapú robusztus szűrő módszert, amely a leggyakoribb érték módszerét is magában foglalja. Eljárásomat mind a mediánszűrés, mind az MFV eljárás által kapott eredményekkel összehasonlítottam.

Amennyiben az adatmátrixok adatpontjainak nagy százaléka kiugró zajjal szennyezett, a zajcsökkentési eljárás jobb eredményt adhat, mint a hagyományos mediánszűrés, ha a potenciálisan nagyobb eséllyel zajnak minősülő pontokat eltávolítjuk a bemeneti adathalmazból, mielőtt a mediánértéket kiszámítanánk. Ugyanezen mátrixok esetében azonban nagy valószínűséggel adódnak olyan adatrészletek, amelyeknél a szűretlen bemeneti adatok jobb eredményt adnak. Ezen két eset közötti automatikus súlyozás adja az alapját szűrési eljárásomnak [S4].

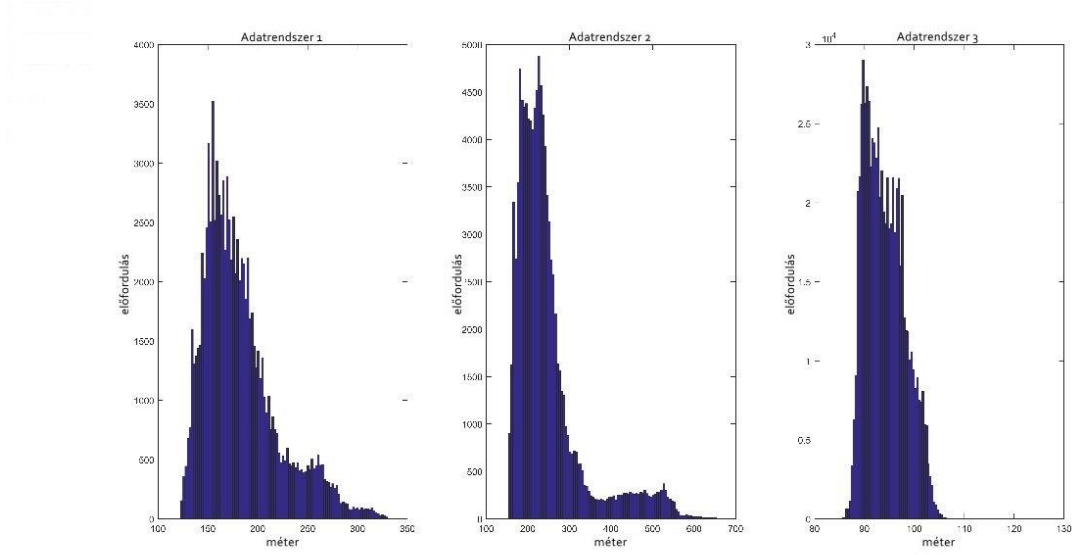
A következő fejezetekben ismertetem a módszer elvét, változóinak kezdeti hangolását (melynek alapja az [S5] publikációban olvasható), majd összehasonlítom az eljárást mind a hagyományos mediánszűréssel, mind a leggyakoribb érték módszerrel, mivel ez utóbbi sokkal hatékonyabb, mint a szokásos módszerek. A fenti megoldások hatékonyságát digitális domborzati adatrendszerek zajcsökkentése esetében vizsgálom, különböző zajszintek mellett. Az összehasonlítás főként három mérőszám alapján történik, a legnagyobb hangsúlyt a L_1 norma-alapú távolságmetrika eredményekre téve. Ezt követően a módszer egy módosított változatát – amely a Steiner-féle MFV-szűrőt tartalmazza központi elemként – is bemutatom, hasonló vizsgálattal.

Az MFV módszer hatékonyságát az adja, hogy a leggyakoribb érték számításához szükséges súlyokat automatikusan szolgáltatja a bemenő adatokat terhelő zaj eloszlásától függetlenül (robusztusság). E miatt a legtöbb esetben különböző mértékű és eloszlású zaj esetén jobbnak bizonyul a hagyományos mediánszűrésnél. Eljárásom jobbnak bizonyul számos esetben a Steiner-féle MFV-nél is, a nem nulla átlagú zajok kezelésében. A módszer módosított változata – amely a Steiner-féle MFV módszert is magában foglalja – a dolgozatban ismertetett alkalmazási területen a nulla átlagú zajok kezelésében elérte ezt.

Mivel az általam javasolt megoldás mindkét változatának megalkotásakor elsődleges cél a digitális domborzati adatok szórt zajának mozgóablakos eltávolítása volt [S6], így a fenti két bevett eljárást is ilyen adatok esetében hasonlítom össze módszeremmel, a vizsgálatot négy különböző adatpont-arányt érintő zajterhelés esetében is bemutatva.

2.2 Feldogozott adatok

Az elemzett adatok három 25 m térbeli felbontású digitális magassági adathalmazból álltak, amelyek az ArcGIS szoftverben Topo to Raster interpolációval hoztam létre az 1:10000 méretarányú EOTR térképlapok szintvonalainak, magassági pontjainak és vízhálózatának digitalizálásával. A három adatrendszer hisztogramja a 8. ábrán látható. Az első adatrendszerben az adatok átlaga 211,99 m, míg a szórás 8,440 m volt. Ugyanezen a statisztikai adatok a második adatrendszerben 191,31 m átlag és 8,752 m szórás voltak, míg a harmadik adatrendszerben az átlag értéke 96,39 m, a szórás pedig 0,457 m volt.



8. ábra. A bemeneti domborzati adatrendszerek hisztogramjai

A fent jellemzett digitális magassági adatrendszerek mindegyike esetében először normális eloszlású zajt adtam az adatmátrixokhoz, olyan szórással, hogy az átlagos zajamplitúdó az adatmátrix átlagának körülbelül 1%-a legyen. Ezt követően – mint kiugró, nem nulla átlagú impulzus-zaj – a pontok 10, 15, 20 és 25 százalékához véletlenszerűen további zajt adtam. Ennek érdekében minden sorhoz egy normális eloszlású zajvektort generáltam, amelynek átlaga megegyezett az adott adatsor átlagával, szórása pedig olyan volt, hogy a zaj átlagos amplitúdója az adatsor átlagának 100%-a körül legyen. Ezután az adott adatsorhoz generált zajvektor elemeit véletlenszerűen szórtam szét a sorban 0,1-0,7 közötti amplitúdó szorzóval (a fentiek szerint az adatpontok 10, 15, 20, ill. 25 százalékán), ami további ~10-70%-os zajt adott az adatokhoz a különböző tesztesetekben (a továbbiakban 0,1-0,7 zajamplitúdó).

2.3 A súlyozott medián (WM) módszer

A módszer segítségével a képmátrixokon végig haladva a mozgóablak minden egyes pozíciójában annak központi elemének korrigált értékét állítottam elő az ablak többi elemének súlyozott átlaga segítségével. Az eljárás ezen változata elsősorban a nem nulla átlagú zajok kiküszöbölésére szolgál (amely például a mérőeszköz hibája, vagy a hosszú távú adatátvitel során jött létre).

A súlyozott átlagot futóablakkal minden egyes adatpontra kiszámítom, az alábbiakban meghatározott két súllyal (w_1, w_2). Ennek érdekében két független ablakszűkítési folyamat történik kezdeti lépésként, a súlyok kiszámítása előtt. Ezeket a szűkített ablakokat az aktuális (5x5 méretű) adatablakból hozom létre, minden ablakpozícióban. Tesztjelleggel 9 (3x3) és 49 (7x7) ablakméretet is kipróbáltam, azonban ezek nem bizonyultak optimálisnak a probléma megoldásához.

Az első ablakszűkítés folyamata a következő. A mozgóablakban lévő elemek értéktartományát két és három egyenlő szélességű tartományra (binre) osztjuk, majd két arányszámot hozunk létre:

- λ_1: a nagyobb és a kisebb elemszámú tartomány közötti elemszámarány a kettő közül,
- λ_2: a legnagyobb és a második legnagyobb tartományok aránya a 3-ból (ismét az elemszám tekintetében).

Ha $\lambda_1 > \lambda_2$, az új halmaz (D) a 2 tartomány legnagyobb elemszámú tartománya, egyébként a 3 tartomány legnagyobb elemszámú tartománya. Ekkor m_s a következő formában definiálható:

$m_s = medián(D)$.

Így magasabb λ értéket az élesebb vágás jeleként használom. Azért választottam a 2 és a 3 értéket a tartományok ideális számaként, mert 4 (vagy több) részre való felosztás esetén egy-egy bin egyes esetekben nem tartalmaz elegendő számú elemet a kezdeti 5x5-ös ablakból ahhoz, hogy elvégezhessem a leírt további lépéseket.

Ezután egy másik (független) második ablakszűkítési folyamat következik, m_{e1} paraméter értékének meghatározására. Ennek érdekében az eredeti mozgóablak elemei érték szerint rendezésre kerülnek, majd két ill. három egyenlő szélességű tartományra kerülnek felosztásra (az értékkészlet alapján).

Például a három részre osztás esetén, ha a rendezett vektor v, és $max(v)$ annak legnagyobb értékű eleme, $min(v)$ pedig a legalacsonyabb, akkor a harmadik tartományban – amelyben a legalacsonyabb értékkel rendelkező értékek vannak – $\big(max(v) - min(v)\big)/3$-nál alacsonyabb értékek lesznek.

Itt kiszámítjuk λ_3 arányt a két halmazra való osztás esetéből. λ_3 értéke $1/n$, ahol n a két tartomány elemszámának összege a legmagasabb értékeket tartalmazó intervallum nélkül (azaz itt valójában a kisebb értékeket tartalmazó intervallum elemszáma – a fenti megfogalmazás az általánosítás okán az előbbi formában írt). Ezután kiszámítjuk λ_4 arányt abból az esetből, amikor három tartományra osztunk. Ennek az aránynak az értéke $1/m$, ahol m a három tartomány elemszámának összege ismét a legmagasabb értékű szelet nélkül.

Amennyiben $\lambda_3 > \lambda_4$, az esetben a kettő közül a legnagyobb értékeket tartalmazó halmazt, ellenkező esetben a három közül a legnagyobb értékeket tartalmazó tartományt választjuk a kiválasztott halmaznak (E). Végül, m_{e1} a kiválasztott új halmaz elemeinek átlaga lesz.

Az előbbihez hasonló érték, m_{e2} ugyanezzel a módszerrel határozható meg, azonban úgy, hogy az eredeti ablakot 3 és 5 részre osztjuk (2 és 3 helyett). Ezzel a lépéssel magasabb osztásszám jelenik meg az eredményben, ha az aktuális ablak értékkészlete ezt lehetővé teszi (azaz amennyiben az új intervallumok elemszáma nem nulla).

Mind m_{e1} mind m_{e2} az ablak értékkészletéből kiválasztott legnagyobb értékek kiválasztott részhalmazán kerül kiszámításra, tehát mindkét érték a maximumokkal áll kapcsolatban.

Fontos különbség m_{ex} és m_s között, hogy míg m_{e1} ill. m_{e2} esetében az átlagot számítjuk, m_s-nél mediánt számítunk, különböző tartományszámokkal. Emellett, míg m_{e1} és m_{e2} esetében a legszűkebb halmazban a legnagyobb értékű elemek vannak, m_s esetében a kiválasztott szűkített

halmaz az adott feltételek melletti legtöbb elemet tartalmazó halmaz (amely nem feltétlenül a legnagyobb értékeket tartalmazza).

Most már kiszámíthatjuk az első súlyt (w_1), az aktuális pont súlyvektorának súlyát a következőképpen:

$$w_1 = (m_s / m_{e1}) * \alpha, \quad (1)$$

ahol α egy skálázási tényező, amely biztosítja, hogy a w_1 értékei ugyanabba az értéktartományba esnek, mint w_2, az alábbiakban leírtak szerint. Ezen skálázási tényező értékét ($\alpha = 1/3$) kísérleti alapon határoztam meg, hogy megfeleljen ennek a célnak. Megjegyzendő, hogy ez adott adatrendszer esetében a fentiek figyelembevételével megváltoztatható, amennyiben w_1 ill. w_2 súlyok aránya a fentiektől nagymértében eltér.

Meghatározhatjuk m_w értékét is, mint az eredeti mozgóablakban lévő értékek mediánját.

A fent leírt értékek felhasználásával három részsúlyt kapunk az alábbiak szerint (mindegyiknek szerepe van w_2 súly értékének meghatározásában).

$$w_a = \frac{1}{max_1} * (m_{e1} - max_1) + 1, \ (2)$$

ahol max_1: $max(m_w, m_{e1})$,

$$m_{as} = |m_{e1} - m_{e2}|, \quad (3)$$

$$w_p = \frac{\beta}{max_2} * (m_{as} - max_2) + 1 , \quad (4)$$

ahol max_2: $max(m_w, m_{e1}, m_{e2})$,

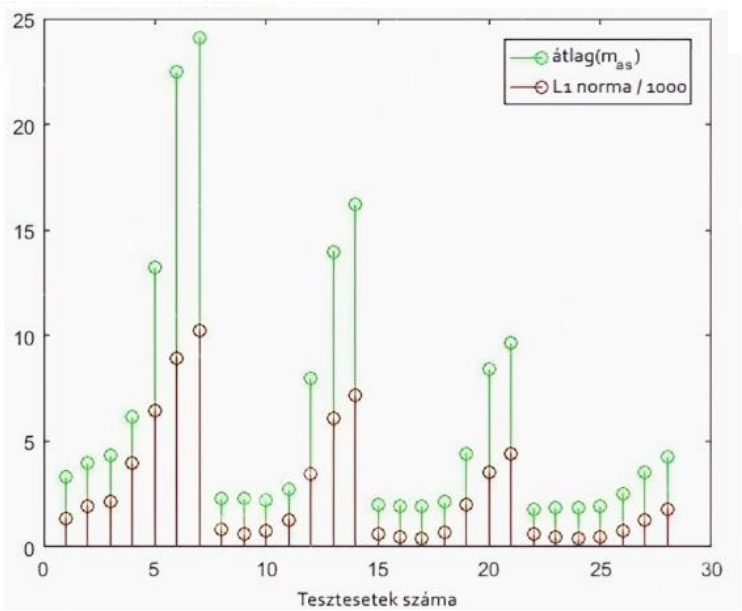
$$w_{p2} = \frac{\gamma}{max_3} * (m_{as} - max_3) + 1 , \quad (5)$$

ahol max_3: $max\big(m_w, átlag(ablak)\big)$.

Az adott részsúlyban (w_a,w_p, w_{p2}) alkalmazott maximális érték kiszámítása minden esetben tartalmazza a teljes ablak mediánját, valamint az átlag (m_{e1}, m_{e2}) vagy a medián (m_s) értékét a szűkített ablakon.

A 4. és az 5. egyenletben kevesebb elem marad ki az eredeti ablakból ($m_{e2} - nek\ m_{as}$ általi tartalmazása okán), így w_p, és w_{p2} részsúlyok kisebb szorzóval számítandók, mint a 2. egyenletben lévő szorzó. Ezen β és γ szorzó konstansok értékei 0,5-ben kerültek meghatározásra. A hangolási eljárás eredményei a 2.6. szakaszban láthatók.

Mind w_p mind w_{p2} súlyok korrekciós szerepet töltenek be. Értékük magas lehet, ha a súlyok 3 és 5 részre osztással kapott részhalmazok átlagai (m_{e1} és m_{e2}) között nagy a különbség. Amint az a 9. ábrán látható, m_{e1}és m_{e2} közötti nagy különbség nagy L_1 norma-hibát eredményez. Tehát, a nagy különbség azt jelzi, hogy az aktuális ablakpozícióban végzett hisztogramműveletek torzíthatják az eredményt, ezért ennek elkerülése érdekében a különbség értékének növekedése növeli w_2 súlyát (azaz a hagyományos, hisztogramműveletek nélküli medián súlyát).



9. ábra. L1 norma és m_{as} érték viszonya

A w_a rész súly 1-től eltérő, magas értéket az esetben vehet fel, ha az eredeti, nem szűkített ablak m_w medián értéke nagyobb, mint a szűkített ablak elemeinek átlaga (m_{e1}). Mivel a szűkített ablak tartalmazza a részhalmazok legnagyobb értékeit, ha az eredeti ablak mediánja és e szűkített ablak elemeinek átlaga közötti különbség kiemelkedő, az azt valószínűsíti, hogy a hisztogramműveletek az aktuális ablakpozícióban torzítják az eredményt. Az előzőekhez hasonlóan, nagy pozitív különbség m_{e1} és max_1 között nagy L_1 norma-hibát eredményez, ezért ennek tükröződnie kell a végső w súlyvektorban vagy w_1 értékének csökkenése, vagy w_2 értékének növelése formájában (azaz a hagyományos medián módszer eredményének súlyának növelésével). Ez utóbbi w_2 súlyban w_a használatával valósul meg. Mivel w_a a három

korrekciós tényező (w_a, w_b, w_c) közül a legfontosabb, a négyzete szerepel w_2 képletében. A nagyobb hatást a súly négyzetre emelésével értem el, mivel a súly maximális értéke – mielőtt egyet hozzáadnánk az értékéhez – 1, így a súly négyzetre emelése még a maximális értékénél sem eredményez szélsőséges súlyértéket.

A +1 w_p és a w_{p2} részsúlyok képleteiben amiatt szerepel, mivel mindegyikben szerepel a maximális érték kivonása, ami számos esetben negatív értéket eredményez, így a konstans a pozitív tartományba való eltolódást biztosítja. w_a képletében az 1 hozzáadásának szerepe az, hogy minimumát egynél nagyobb értékre tolja (annak érdekében, hogy növelni lehessen a w_2 súlyát, amely függ ennek értékétől).

Végül a következő két súlyt (w_b, és w_c) w_p és w_{p2} felhasználásával állíthatjuk elő:

$$w_b = 1 + \frac{w_p}{2}. \quad (6)$$

Mivel w_c részsúly maximális értéke nem a különböző szűkített ablakok, hanem az eredeti ablak mediánjának vagy átlagértékének függvénye, ezt a részleges súlyt kisebb konstanssal vesszük figyelembe:

$$w_c = 0.5 - \frac{w_{p2}}{2}. \quad (7)$$

A fent meghatározott összetevőkkel w_2 súly a következő formát ölti:

$$w_2 = {w_a}^2 * w_b * w_c. \quad (8)$$

Ezen a ponton már ismerjük az aktuális adatpont w súlyvektorát:

$$w = [w_1\ w_2]^T. \quad (9)$$

Megjegyzendő, hogy w súlyvektorban (9. egyenlet) $w_1$és w_2 súlyok egyrészt az aktuális adatablak mediánjára (m_w), másrészt ugyanezen ablak redukált halmazának mediánjára (m_s) hatnak, w_1 által az utóbbit, w_2 által az előbbit a következőképpen súlyozva (például az adatmátrix k-adik elemére):

$$res_{WM_k} = (w_1 * m_s + w_2 * m_w)/(w_1 + w_2). \quad (10)$$

A fent leírtak szerint a szűkített ablak mediánja (m_s), és az eredeti ablak mediánja (m_w) minden egyes ablakpozícióban súlyozásra kerül az aktuális pont végeredményéhez. Ehhez m_s w_1 súlya fent ismertetett $(m_s\ /m_{e1}) * \alpha$ (azaz a mediánt elosztjuk a szűkített ablak maximális értékeinek

átlagával). Ha ez az arány például a zaj miatt alacsony az adott mozgóablak-pozícióra (vagyis nagy értékű a maximumok m_{e1} átlaga), akkor $m_s$súlyának arányosan alacsonynak kell lennie, egyébként a kiugró maximumok magas értéke negatívan hatna a végeredményre. Ilyen esetekben m_w a súlya arányosan nagy lesz – nem csupán m_s kis súlya miatt, hanem azért is, mert $m_w$súlyai w_a,w_b, w_c, amelyek mindegyike tartalmazza m_{e1}vagy m_{e2} értékeket.

2.4 Leggyakoribb érték módszer

A számtani átlagnál sokkal megbízhatóbb statisztikai jellemzőt, a súlyozott átlagot úgy kapjuk, hogy viszonylag kis súlyt (w_k) rendelünk az adatok többségétől (X_k) távol eső pontokhoz és nagyobb súlyt (w_k) a legnagyobb adatsűrűségű pontokhoz (11. egyenlet)

$$M = \sum_{k=1}^{N} X_k w_k \left[\sum_{k=1}^{N} w_k\right]^{-1} \text{ (k=1,2,...,N) .} \qquad (11)$$

A k-adik súlyt Steiner [46] a következőképpen választotta ki:

$$w_k = \varepsilon^2/[\varepsilon^2 + (X_k - M)^2] \, . \qquad (12)$$

A fentiekben, N az adatok száma, ε pedig a dihézió, az ismeretlen skalár paraméter, amelyet optimalizációs eljárás keretében határozunk meg. Ha ε nagy, akkor minden adat közel azonos súlyt kap, és a kiugró értékek elrontják az értékbecslést, ha pedig ε túl kicsi, akkor ügyelni kell arra, hogy egyes adatokat ne hagyjunk figyelmen kívül.

A (11) egyenlet által meghatározott leggyakoribb értéknek (M) nevezett súlyozott átlagot előre ismernünk kell ahhoz, hogy a környezetéhez maximális, tőle távolabb pedig egyre kisebb súlyokat lehessen hozzárendelni. Ezért ez az eljárás iteratív algoritmust igényel, amelyben az M és az ε együttesen kerül meghatározásra. Az első iterációs lépésben a dihézió a mintatérből a következő képlet segítségével becsülhető meg:

$$\varepsilon_1 = \left(\sqrt{\frac{3}{2}}\right)[max(X_k) - min(X_k)] \, , \qquad (13)$$

míg a M_1 esetében a kiindulási értéket lehetőleg a minta átlagának vagy mediánjának választjuk. Kutatásaim során a kettő közül a kieső adatokra kevésbé érzékeny mérőszámot, a mediánértéket használtam.

A következő iterációs lépésekben, M és ε az alábbiak szerint származtathatók egymásból:

$$\varepsilon_{j+1}^2 = \frac{3\sum_{k=1}^{N} \frac{(X_k - M_j)^2}{\left[\varepsilon_j^2 + (X_k - M_j)^2\right]^2}}{\sum_{k=1}^{N} \frac{1}{\left[\varepsilon_j^2 + (X_k - M_j)^2\right]^2}} \leftrightarrow M_{j+1} = \frac{\sum_{k=1}^{N} \frac{\varepsilon_{j+1}^2}{\varepsilon_{j+1}^2 + (X_k - M_j)^2} X_k}{\sum_{k=1}^{N} \frac{\varepsilon_{j+1}^2}{\varepsilon_{j+1}^2 + (X_k - M_j)^2}}. \qquad (14)$$

2.5 Az eredmények minőségellenőrzése

A fent ismertetett, saját (súlyozott medián), illetve összehasonlításul alkalmazott egyéb (MFV, medián) szűrési módszerek eredményeinek összehasonlításához a következő mérőszámokat használom dolgozatomban. Az RMSE (Root Mean Square Error -négyzetes hiba) értékének kiszámítása az MFV módszer és a medián szűrés esetében (ahol *inp* a zajmentes adatmátrix):

$$RMSE_{St} = \sqrt{\frac{\sum_{i=1}^{N}(res_{St_i} - inp_i)^2}{N}}, \quad (15)$$

ahol res_{St}*: a* Steiner-féle MFV-vel korrigált mátrix ,

$$RMSE_{Med} = \sqrt{\frac{\sum_{i=1}^{N}(res_{Med_i} - inp_i)^2}{N}}, \quad (16)$$

ahol res_{Med}*:* medián módszerrel korrigált mátrix,

$$RMSE_{WM} = \sqrt{\frac{\sum_{i=1}^{N}(res_{WM_i} - inp_i)^2}{N}}, \quad (17)$$

ahol res_{WM}*:* súlyozott mediánnal korrigált mátrix.

A szórás a három eljárás tekintetében:

$$Std_{St} = \overline{std(res_{St} - inp)}, \quad (18)$$

$$Std_{Med} = \overline{std(res_{Med} - inp)}, \quad (19)$$

$$Std_{WM} = \overline{std(res_{WM} - inp)}, \quad (20)$$

ahol res_{St}, res_{Med}, res_{WM} *inp*: a fentiek szerint.

AzL_1 norma alapú mérőszámok a következők:

$$L_{1St} = \|res_{St} - inp\|_1, \quad (21)$$

$$L_{1Med} = \|res_{Med} - inp\|_1, \quad (22)$$

$$L_{1WM} = \|res_{WM} - inp\|_1, \quad (23)$$

ahol res_{St}, res_{Med}, res_{WM}, *inp*: a fentiek szerint.

2.6 Konstansok beállítása

Megvizsgáltam a részsúlyok képleteiben (4), (5) alkalmazott konstansok ideális értékét is, így elsőként a β konstanst. Ennek kezdeti 0,5 értéke környezetében, néhány véletlenszerűen kiválasztott tesztesetben azt vizsgáltam, hogy hogyan alakul az L_1 norma-távolság értéke a zajmentes mátrix és a súlyozott medián-korrigált mátrix között a konstans különböző értékeivel. A norma értéke minden ilyen esetben monoton módon csökkent a konstans 0,5 értékéig, azt követően pedig nőtt, amint az az első táblázatban lévő példában látható (10%-os zajszint, 0,3 zajamplitúdó esetén):

1. táblázat Példa β korrekciós állandó konstansának beállítására

β	2	1.5	1	0.5	0.25	0.125	0.0625	0.03125
L_1 norma távolság	843.12	822.32	801.63	762.82	794.51	956.47	1720.47	2393.27

Hasonlóképpen vizsgáltam a w_{p2} részsúlyban szereplő γ-t is, ami jellegében az előbbi konstans esetében tapasztaltakkal megegyező eredményeket mutatott, és itt is a 0,5-ös érték bizonyult ideálisnak, amint az a 2. táblázat példáján látható (ugyanazoknál a paramétereknél, mint az előző esetben):

2. táblázat Példa a számításhoz használt γ korrekciós állandó beállítására

γ	2	1.5	1	0.5	0.25	0.125	0.0625	0.03125
L_1 norma távolság	735.13	721.78	698.67	654.31	668.807	668.94	676.11	685.71

Megjegyzendő, hogy a fenti konstansok globális optimumának beállítása ez idáig nem volt része a vizsgálatoknak.

2.7 Összehasonlító vizsgálatok

A szűrési eljárások összehasonlító vizsgálataként minden esetben egy 5x5-ös ablakot futtattam végig az adatmátrixokon, és mind a Steiner-MFV módszer szűrőként használva [13], mind az általam megalkotott súlyozott medián módszer mindig az ablak középső elemének értékét korrigálta, a korrekcióhoz az ablak összes elemét bemenetként használva.

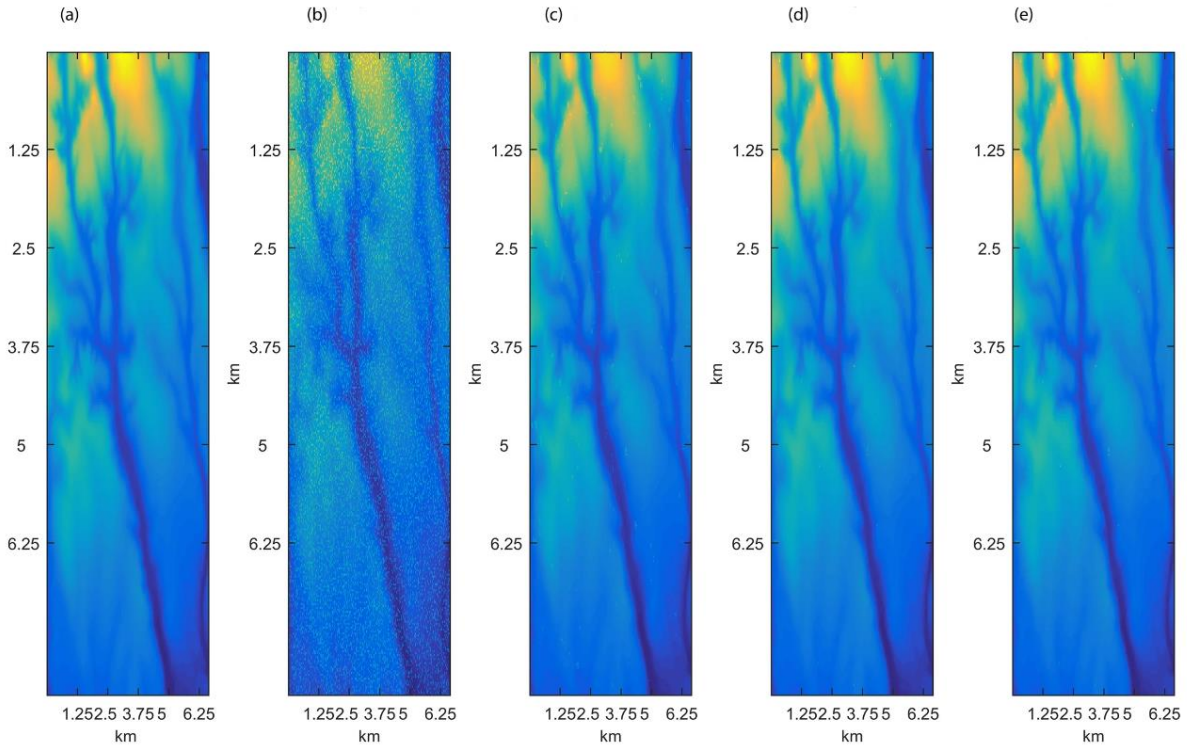
A Steiner-szűrő iterációinak száma minden esetben 20 (mivel próbafuttatások során 20 iteráció felett már nem történt változás), a kezdeti súly pedig az ablakelemek mediánja volt.

Az eredményeket minden esetben összehasonlítottam a klasszikus mediánszűrővel kapott eredményekkel is. A mediánszűrő ugyanazokon a zajos adatmátrixokon futott ugyanazzal az ablakmérettel, mint a Steiner-módszer és a súlyozott medián módszer.

A 3. táblázatban látható példa a metrikák első adatsorozatra vonatkozó értékeit mutatja, az adatpontokon 0,3-as zajamplitúdójú 15%-os zaj esetében.

3. táblázat Példa az összehasonlító eredményekre (10%-os Gauss eloszlású? kiegészítő kiugró zaj az adatpontok 15%-ánál)

Minőségi paraméter	Érték
L_{1WM}	538.98
L_{1St}	630.41
L_{1Med}	711.19
$RMSE_{WM}$	1.42
$RMSE_{St}$	2.02
$RMSE_{Med}$	2.13
Std_{WM}	1.19
Std_{St}	1.75
Std_{Med}	1.58



10. ábra Szűrési eredmények vizuális reprezentációja

Egy másik példa látható a 10. ábrán, ahol ismét zajjal terhelt az adatpontok 15%-a, ez esetben 0,5-ös zaj amplitúdó szorzóval. 10. ábra: (a) az eredeti adatok, (b) a zajos adatok, (c) a Steiner-féle MFV módszer eredménye, (d) a súlyozott medián módszer eredménye, (e) a klasszikus medián módszer eredménye.

A 4. táblázat egyrészt L_1 normák segítségével mutatja a zajmentes bemeneti adatmátrixtól való távolságot, másrészt a normák arányát is a súlyozott medián módszer (L_{1WM}) és a Steiner-módszer (L_{1St}) esetében, a pontok 25%-ának zajjal való szennyezése mellett, különböző zajamplitúdók (0,1,...,0,7) függvényében az első adatsoron. Az értékek azt mutatják, hogy két esetben a Steiner-módszer adott jobb eredményt, mintegy 6%-kal, a többi esetben pedig a súlyozott medián módszer bizonyult jobbnak. Ez utóbbi átlagosan 6,3%-kal jobb eredményt produkált (mivel az L_{1WM}/L_{1St} arányok átlaga 0,937).

4. táblázat L_1 norma értékek 25%-os zajarány mellett

Zaj ampl. szorzó	0.1	0.2	0.3	0.4	0.5	0.6	0.7
L_{1WM}	540.22	425.44	420.39	471.26	615.92	787.36	515.82
L_{1St}	597.90	457.46	536.20	541.64	580.61	738.81	545.17
L_{1WM}/L_{1St}	0.90	0.92	0.784	0.87	1.06	1.06	0.94

Az 5. táblázat a súlyozott medián eljárás és a standard mediánszűrés eredményeit mutatja azonos zajszint mellett. Ebben az esetben a súlyozott medián eljárás átlagosan 26,4%-kal jobbnak bizonyult az adathalmazon.

5. táblázat L_1 norma értékek 25%-os zajarány mellett

Zaj ampl. szorzó	0.1	0.2	0.3	0.4	0.5	0.6	0.7
L_{1WM}	540.22	425.44	420.39	471.26	615.92	787.36	515.82
L_{1Med}	604.39	582.33	725.92	692.10	827.83	970.98	723.63
L_{1WM}/L_{1Med}	0.89	0.73	0.57	0.68	0.74	0.81	0.71

A 6. és 7. táblázat a korábbiakkal megegyező összehasonlítást mutatja, a pontok 20%-át érintő zajra vonatkozó L_1 norma értékével. Ebben az esetben megfigyelhetők olyan zajamplitúdó értékek, ahol a Steiner-módszer kisebb távolságot ad a zajmentes mátrixtól, mint a súlyozott

medián eljárás, a normál medián eljárás azonban egyik esetben sem tudta ezt elérni. A súlyozott módszer átlagosan 23%-kal jobb az utóbbinál.

6. táblázat L1 normaértékek 20%-os zajarány mellett

Zaj ampl. szorzó	0.1	0.2	0.3	0.4	0.5	0.6	0.7
L_{1wm}	566.65	437.97	440.14	485.61	708.77	646.18	759.37
L_{1st}	545.34	468.96	469.76	608.38	1002.99	866.07	841.60
L_{1wm}/ L_{1st}	1.03	0.93	0.93	0.79	0.70	0.74	0.90

7. táblázat L1 normaértékek 20%-os zajarány mellett

Zaj ampl. szorzó	0.1	0.2	0.3	0.4	0.5	0.6	0.7
L_{1wm}	566.65	437.97	440.14	485.61	708.77	646.18	759.37
L_{1med}	612.87	627.87	612.76	741.06	901.39	792.62	979.72
L_{1wm}/ L_{1med}	0.92	0.69	0.71	0.65	0.78	0.81	0.77

A 8. és 9. táblázat a 15%-os zajos pontok esetét mutatja.

8. táblázat L1 normaértékek 15%-os zajarány mellett

Zaj ampl. szorzó	0.1	0.2	0.3	0.4	0.5	0.6	0.7
L_{1wm}	542.97	433.71	538.98	464.25	493.81	675.11	624.08
L_{1st}	519.85	473.9043	630.41	490.20	561.66	732.78	644.35
L_{1wm}/ L_{1st}	1.04	0.91	0.85	0.94	0.87	0.92	0.96

A 8. táblázat eredményei, amelyek a Steiner-módszert és a súlyozott medián eljárást hasonlítják össze, azt mutatják, hogy 0,1-es zajamplitúdó-szorzó esetén a Steiner-módszer használata előnyösebb.

9. táblázat L1 normaértékek 15%-os zajarány mellett

Zaj ampl. szorzó	0.1	0.2	0.3	0.4	0.5	0.6	0.7
L_{1wm}	542.97	433.71	538.9896	464.25	493.81	675.11	624.08
L_{1med}	604.28	614.34	711.1921	639.07	672.32	846.64	695.19
L_{1wm}/ L_{1med}	0.89	0.70	0.75	0.72	0.73	0.79	0.89

A súlyozott medián eljárás átlagosan 21,2%-kal jobb teljesítményt nyújt a súlyozatlan medián eljárásnál az első adatsorra 15%-ban zajos pontok esetén (9. táblázat). A 10. és 11. táblázat a három eljárás adattávolságait mutatja be L_1 norma szerint arra az esetre, amikor az adatok 10%-a volt zajjal szennyezett.

10. táblázat L1 normaértékek 10%-os zajarány mellett

Zaj ampl. szorzó	0.1	0.2	0.3	0.4	0.5	0.6	0.7
L_{1wm}	521.53	483.07	543.86	484.36	779.51	727.48	724.20
L_{1st}	507.13	482.52	542.96	527.25	652.87	939.20	790.58
L_{1wm}/ L_{1st}	1.02	1.00	1.00	0.91	1.19	0.77	0.91

Ebben az esetben (zajjal a pontok 10%-án) a Steiner-módszer a zajamplitúdók többsége esetén – 4 esetben, a négy esetben átlagosan 5,6%-kal – felülmúlja a súlyozott medián módszerét, a fennmaradó három esetben pedig a WM-módszer 13%-kal teljesít jobban az adathalmazon.

11. táblázat L1 normaértékek 10%-os zajarány mellett

Zaj ampl. szorzó	0.1	0.2	0.3	0.4	0.5	0.6	0.7
L_{1wm}	521.53	483.07	543.86	484.36	779.51	727.48	724.20
L_{1med}	605.38	637.25	746.44	645.17	871.46	1038.26	865.18
L_{1wm}/ L_{1med}	0.86	0.75	0.72	0.75	0.89	0.70	0.83

A 12. táblázat a három eljárás esetében számított RMSE értékeket és azok arányát mutatja a baloldalon 25%-os, a jobboldalon 20%-os zajterhelés esetén, mindkét esetben különböző zajamplitúdó-szorzók (0,1,...,0,7) esetén.

12. táblázat RMSE-értékek 25%-os és 20%-os zajarány mellett

Zaj arány	25%							20%						
Zaj ampl. szorzó	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.1	0.2	0.3	0.4	0.5	0.6	0.7
$RMSE_{WM}$	1.54	1.47	1.55	2.04	2.50	2.73	2.25	1.58	1.51	1.71	1.63	2.14	2.52	2.74
$RMSE_{Med}$	1.79	1.95	2.16	2.50	2.84	3.07	2.64	1.84	2.03	2.30	2.17	2.58	2.95	3.10
$RMSE_{St}$	1.60	1.74	2.07	2.73	3.29	3.79	3.76	1.63	1.82	2.21	2.50	3.09	3.61	4.24
$RMSE_{WM}$ /$RMSE_{Med}$	0.86	0.75	0.72	0.82	0.88	0.89	0.85	0.86	0.75	0.74	0.75	0.83	0.85	0.88
$RMSE_{WM}$ /$RMSE_{St}$	0.96	0.84	0.75	0.75	0.76	0.72	0.60	0.97	0.83	0.77	0.65	0.69	0.70	0.65

A 13. táblázat az RMSE-értékeket mutatja az előzőekhez hasonlóan, itt az adatpontok 15%-ának és 10%-ának zajszintje esetén.

13. táblázat RMSE-értékek 15%-os és 10%-os zajarány mellett

Zaj arány	15%							10%						
Zaj ampl. szorzó	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.1	0.2	0.3	0.4	0.5	0.6	0.7
$RMSE_{wm}$	1.59	1.55	1.42	1.70	1.83	2.34	2.55	1.60	1.46	1.75	1.50	2.12	2.02	2.72
$RMSE_{med}$	1.83	2.04	2.13	2.22	2.36	2.74	2.92	1.86	2.00	2.29	2.10	2.56	2.60	3.08
$RMSE_{St}$	1.62	1.83	2.02	2.55	2.97	3.72	3.81	1.66	1.84	2.13	2.36	3.01	3.32	4.23
$RMSE_{wm}$ /$RMSE_{med}$	0.87	0.76	0.67	0.77	0.77	0.85	0.87	0.86	0.73	0.77	0.72	0.83	0.78	0.88
$RMSE_{wm}$ /$RMSE_{St}$	0.98	0.85	0.70	0.67	0.61	0.63	0.67	0.96	0.80	0.82	0.64	0.71	0.61	0.64

Látható, hogy a súlyozott medián módszer rosszabbul teljesít a legnagyobb zajamplitúdó-szorzóknál (0,7), azonban ez a többi módszerre is igaz, így az azokhoz viszonyított aránya nem romlik. Az MFV módszerrel kapott RMSE a legkisebb zajamplitúdó-szorzó (0,1) esetén áll a legközelebb a súlyozott medián módszerrel kapott RMSE-hez. A súlyozott medián módszer RMSE értéke átlagosan 79,5%-a a hagyományos medián módszerrel kapott értéknek.

14. táblázat A két medián módszer szórásértékei 10%-os zajszint mellett

Zaj ampl. szorzó.	0.1	0.2	0.3	0.4	0.5	0.6	0.7
Std_{WM}	1.41	1.25	1.37	1.20	1.43	1.37	1.62
Std_{Med}	1.42	1.51	1.70	1.56	1.76	1.74	1.96
Std_{WM} / Std_{Med}	0.99	0.82	0.80	0.77	0.81	0.79	0.83

Mivel a súlyozott medián eljárás akkor bizonyult a legkevésbé hatékonynak, amikor a fenti tesztek során az adatpontok 10%-a volt zajjal szennyezett, ebben az esetben a szórásokat is megvizsgáltam. Ennek példáját a standard medián módszerrel összehasonlítva a 14. táblázat mutatja be.

15. táblázat L1 norma arányok a két vizsgálati adathalmazra együttesen

Zaj mérték százalék	$min(L_{1WM}/L_{1St})$	$max(L_{1WM}/L_{1St})$	$avg(L_{1WM}/L_{1St})$
25%	0.81	1.07	0.93
20%	0.81	0.99	0.91
15%	0.83	0.98	0.90
10%	0.80	1.07	0.93

A 15. táblázat az első és a második adatsor eredményeinek átlagát mutatja. Az L_1 normás adattávolsági arányok minimális, maximális és átlagos értékei a különböző zajszintek függvényében a Steiner-módszer és a súlyozott medián módszer esetében láthatók. A minimumok átlaga (azaz azon esetek átlaga, ahol a legnagyobb különbség a súlyozott medián módszer javára) 0,82, azaz ezekben az esetekben a módszer 18%-kal jobbnak bizonyult. A maximumok átlaga 1,032 volt, azaz az ellenkező esetben a Steiner-módszer átlagosan 3,2%-kal jobb a két adatrendszerre átlagolva.

16. táblázat L1 norma arányok a két adatkészletre együttesen

Zaj mérték százalék	$min(L_{1WM}/L$	$max(L_{1WM}/L$	$avg(L_{1WM}/L_1$
25%	0.63	0.86	0.74
20%	0.67	0.86	0.77
15%	0.73	0.84	0.78
10%	0.72	0.85	0.77

A 16. táblázat a L_1 normák eredményeit ugyanazon módszer szerint csoportosítva mutatja, ez esetben a két medián eljárásra számítva. A minimumok átlaga 0,692, tehát azokban az esetekben, ahol a súlyozott medián eljárás a jobb, ez a módszer több mint 30%-kal múlja felül a másikat. A maximumok átlaga 0,86, tehát a súlyozott medián eljárás még a legrosszabb esetekben is átlagosan 14%-kal jobb, mint a standard mediánszűrés a két adatsor esetében együttesen.

Egy kevésbé részletes összehasonlítást végeztem a harmadik adatrendszeren (csak az L_1 normák arányait vizsgálva). Ez a vizsgálat hasonló jellemzőket mutatott mint az előzőek, azonban a Steiner-módszer több esetben bizonyult a legjobbnak a három közül, mint korábban. A 10% és 15%-os arányban zajos adatpontok esetében (mindkettő 7 különböző zajamplitúdóval vizsgálva, mint korábban) a Steiner-módszer 14 esetből 11 esetben bizonyult jobbnak a súlyozott medián módszerrel szemben, átlagosan 14,6%-kal (L_1 norma-arányokat tekintve). A 20, ill. 25%-os esetekben a súlyozott medián módszer jobb eredményt adott, 14-ből 14 esetben, átlagosan 19,04%-kal.

A két medián módszer összehasonlítása során az adathalmazon (ismét a L_1normák arányaival), a 28 esetből 27 esetben a súlyozott medián módszer bizonyult jobbnak, átlagosan 14,32%-kal.

2.8 Zéró átlagértékű zajok kezelése

Az általam kialakított módszer korábban bemutatott változatát főként nem nulla átlagértékű zajokra alkottam meg. A következőkben a módszer egy második, szintén általam módosított változatát mutatom be, amelynek célja elsősorban a nulla átlagú normális eloszlású zajok kezelése. A módszer ezen változata a Steiner-féle MFV-értékeket hívja segítségül és használja fel az adott adatablak aktuális központi elemének korrekciójára, a korrekció részeként.

2.8.1 Zajgenerálás

Ami a zajgenerálási folyamatot illeti, az első lépésben az adatmátrixhoz egy általános nulla átlagú zajt adtam. Ennek érdekében minden adatsorhoz sztenderd normális eloszlású zajt generáltam. Annak érdekében, hogy a vizsgált pontok adott százalékához kiugró zajt adjunk, a pontok 20, 15, 10 és 5 százalékához véletlenszerűen további nulla átlagú normális eloszlású zajt adtam, minden ilyen esetben 0,1-0,7 amplitúdójú szorzóval (mint a módszer korábban bemutatott változatában). A zaj szórásának rendre az aktuális adatsor átlagát adtam meg.

2.8.2 A súlyozott medián módszer módosított változata

A módszer korábban bemutatott változatához hasonlóan az első lépések hisztogramszűrések.

Először egy mediánt, majd egy átlagértéket generáltam a szűrt ablakokból, itt azonban mindkettőnél kétlépcsős szűréssel.

Az $m_s$mediánérték előállításához a hisztogram alapú szűrés a következő. Értékkészlet alapján az aktuális adatablak elemeit két és három egyenlő tartományszélességű tartományra (binre) osztjuk, majd két arányszámot generálunk:

- λ_1: a legnagyobb és a második legnagyobb tartományok aránya 2-ből,
- λ_2: a legnagyobb és a második legnagyobb tartományok aránya a 3-ból (mindkét esetben az elemszám tekintetében).

Ha $\lambda_1 > \lambda_2$, az új halmaz (D) a 2 tartomány közül a nagyobb elemszámú tartomány, ellenkező esetben ugyanez a 3 tartomány közül a legnagyobb elemszámú tartománya lesz. Végül m_s értéke a kiválasztott halmaz mediánja lesz: $medián(D)$.

Az m_{e3} értékének meghatározásához egy második szűkítési eljárásra kerül sor. Ebben az eljárásban az ablakelemek először érték szerint rendezésre kerülnek, majd az értékkészlet alapján három egyenlő szélességű tartományra osztódnak. Itt kiszámítjuk λ_3 arányt, mint az ablak elemszámának és a legnagyobb értékű harmad nélküli harmadok elemszám összegének arányát. Ezután kiszámítjuk λ_4-et, amelynek értéke a legnagyobb és a második legnagyobb elemszámú ablakok elemszámának aránya lesz. Tehát mind λ_3-at, mind λ_4-et másképp számolom, mint a módszer előző változatában.

Amennyiben $\lambda_3 > \lambda_4$: a legnagyobb értékeket tartalmazó harmadot, ellenkező esetben a legnagyobb elemszámmal rendelkező tartományt vesszük a kiválasztott halmaznak (E). Így a magasabb λ érték az élesebb vágású esetet reprezentálja.

Ezután vesszük ezt a csonka E halmazt, és az értékeit szintén tartományokra osztjuk. Ezek szélességét a Scott-szabály [41] [42] segítségével határozom meg:

$$3.5 * std\,(E)/numel\left(E^{1/3}\right). \qquad (24)$$

Meg kell határoznunk a tartományok számát is, hogy az összes értéket eloszthassuk bennük (ami egy triviális lépés, mivel ezek szélessége és az adatértékek már ismertek). Kiválasztjuk a legnagyobb elemszámmal rendelkező tartományt, és m_{e3} ennek elemeinek átlaga lesz.

Mivel $m_s$és m_{e3} értékei ezen a ponton rendelkezésre állnak, a tényleges ablak középső elemét helyettesíthetjük m_s-sel (ezzel kialakítva $w_{r_{ms}}$ módosított ablakot), és hasonlóképpen m_{e3}-mal

(ezzel $w_{r_{me}}$-t megalkotva). Végül legyen $w_{r_{St}}$ az ablak, amelynek közepén az MFV módszer eredménye van.

A következő lépésben egyenként összekapcsoljuk $w_{r_{ms}}$, $w_{r_{me}}$ és $w_{r_{St}}$ módosított ablakokat az eredeti (zajos) aktuális ablakkal, ezáltal w_{u1}, w_{u2}, w_{u3} ablakokat képezve.

Így már három gradiens mértékegységet számolhatunk ki a következő módon:

$$G(x,y) = \sqrt{\frac{\frac{\partial w_u}{\partial x} * \partial w_u}{\partial x} + \frac{\frac{\partial w_u}{\partial y} * \partial w_u}{\partial y}}, \qquad (25)$$

$$\mathrm{g} = \frac{1}{|X| * |Y|} \sum_{x \in X} \sum_{y \in Y} G(\mathrm{x},\mathrm{y}). \qquad (26)$$

Nevezzük g_1-nek g értékét akkor, amikor a $w_{r_{St}}$ –t használjuk a számításban, g_2-nek pedig g értékét abban az esetben, ha $w_{r_{ms}}$ -sel számolunk, és g_3-nak $w_{r_{me}}$ esetében.

Ha $g1, g2, g3$ minimuma g_1 (vagyis a Steiner-módszerrel módosított ablak mátrixa a leginkább homogén az eredeti ablakmátrixszal), akkor a tényleges $w = [w_1 w_2]$ súlyvektorban a w_1 értékéül 0-t, w_2 értékéül 1-et adjuk. Így ebben az esetben csak az MFV módszer eredménye fog számítani az adott adatablak korrekciójában. Ha $(g1, g2, g3)$ miniuma g_2, akkor w_1 0,15 és w_2 0,85 értékű lesz. Végül amennyiben $(g1, g2, g3)$ minimuma g_3, akkor w_1 0,4 és w_2 0,6 (vagy ehhez közeli) értékválasztása indokolt tapasztalati alapon. Ebben az esetben tehát az MFV módszer eredményének súlya 0,6 lesz az adott ablakra. Amint láthatjuk, minden esetben súlyozva vannak a Steiner MFV módszer eredményei, és növeljük vagy csökkentjük annak súlyát az aktuális adatablak központi elemének korrekciójában.

Hasonlóan a módszer előző változatához, rossz eredményeket kapunk, ha m_{as}, azaz $|m_{e1} - m_{e2}|$ értéke nagy. Ahhoz, hogy ezt kezelni tudjuk, itt mind m_{e1}-et, mind m_{e2}-t ki kell számítanunk (ugyanúgy, mint az előző változatban), és ha a különbség nagyobb, mint a nyers adatok átlagának néhány %-a (ennek ideális értéke a vizsgált adatsorokra 2), az esetben w_1 értékének 0-nak kell, és a $w_2$értékének 1-nek kell lennie. Az $|m_{e1} - m_{e2}|$ különbség nagy értéke ugyanis, mint a módszer előbbi változatában, ez esetben is azt jelzi, hogy az aktuális ablakpozícióban végzett hisztogramműveletek torzíthatják az eredményt (mivel túl nagy eltéréseket mutatnak annak különböző változatai).

2.8.3 A Súlyozott medián (WM) szűrési eljárás módosított változatának eredményei

A 17. táblázatban az első adathalmazra vonatkozó L_1normaarányok láthatók, a WM módszer eredményeit mind az MFV, mind az eredeti medián módszer eredményével összehasonlítva. Az előbbi esetben a WM módszer a 28 tesztesetből 26 esetben teljesített jobban. Ezekben az esetekben a L_1 normaránya 0,985 volt, tehát a módszer használata átlagosan 1,4%-kal kisebb L_1 norma értéket eredményezett. A fennmaradó két esetben az arány átlaga 1,005 volt, tehát a WM módszer ebben a két esetben 0,5%-kal rosszabbul teljesített. A legjobb L_1 normaarány értéke 0,97 volt, vagyis ekkor a WM-módszer 3%-kal jobb értéket adott a zajcsökkentés során.

Ami az eredeti medián módszerrel való összehasonlítást illeti, a WM módszer mind a 28 esetben jobban teljesített, átlagosan 16,4%-kal (0,836 átlagos L_1 norma-arány érték). Itt a legjobb eredmény 29,5%-os javulás volt (0,705 L_1 norma-arány).

17. táblázat L1 norma arányok az első adatsoron

	25% zaj kitettség							20% zaj kitettség						
L_{1WM}/L_{1St}	0.991	0.987	0.996	1.004	0.991	0.997	0.993	0.978	1.005	0.998	0.975	0.990	0.994	0.997
L_{1WM}/L_{1Med}	0.881	0.847	0.781	0.798	0.777	0.704	0.760	0.903	0.846	0.822	0.795	0.818	0.802	0.807
	15% zaj kitettség							10% zaj kitettség						
L_{1WM}/L_{1St}	0.987	0.982	0.993	0.978	0.991	0.988	0.978	0.979	0.975	0.979	0.972	0.983	0.982	0.970
L_{1WM}/L_{1Med}	0.917	0.879	0.816	0.835	0.861	0.844	0.840	0.922	0.903	0.813	0.833	0.870	0.852	0.878

A 18. táblázat az előzővel azonos szerkezetű formában mutatja az eredményeket, ez esetben a második adatsoron. A Steiner-féle MFV-vel összehasonlítva a WM-módszer jobb eredményeket adott az alábbiak szerint L_1 normát tekintve 23 esetben (átlagosan 1,25%-kal), a fennmaradó 5 esetben pedig átlagosan 0,21%-kal rosszabbul teljesített. A legnagyobb különbség a WM módszer javára 3,12% volt.

A másik medián módszerrel összehasonlítva a WM minden esetben jobban teljesített (átlagosan 16%-kal, maximum 29,4%-kal).

18. táblázat L1 normaarányok a második adatsoron

	25% Zaj kitettség							20% Zaj kitettség						
L_{1WM}/L_{1St}	0.987	0.979	0.994	0.989	0.968	0.991	0.987	0.979	0.994	0.990	0.982	1.000	0.987	1.001
L_{1WM}/L_{1Med}	0.862	0.840	0.897	0.755	0.786	0.775	0.706	0.889	0.843	0.858	0.768	0.793	0.815	0.748
	15% Zaj kitettség							10% Zaj kitettség						

L_{1WM}/L_{1St}	0.995	0.993	0.983	0.987	0.993	0.987	1.000	1.005	0.995	0.983	0.979	0.990	0.989	0.994
L_{1WM}/L_{1Med}	0.919	0.891	0.855	0.809	0.799	0.868	0.857	0.913	0.902	0.884	0.871	0.890	0.829	0.860

A 19. táblázatban a harmadik adatrendszer L_1 norma-arányai láthatók. A WM és az MFV módszer eredményeit összehasonlítva a 28 esetből 23 esetben az előbbi jobban teljesített (átlagosan 1%-kal, maximum 2,4%-kal), az MFV módszer pedig 5 esetben múlta azt felül (átlagosan 1,4%-kal).

19. táblázat L1 normaarányok a harmadik adatsoron

	25% Zaj kitettség							20% Zaj kitettség						
L_{1WM}/L_{1St}	1.000	0.994	1.003	0.989	1.001	0.999	0.990	0.995	0.997	0.995	0.995	0.994	0.982	0.991
L_{1WM}/L_{1Med}	0.904	0.852	0.798	0.939	0.876	0.941	0.882	0.900	0.915	0.815	0.836	0.815	0.856	0.858
	15% Zaj kitettség							10% Zaj kitettség						
L_{1WM}/L_{1St}	0.995	0.991	0.986	0.984	0.996	0.990	0.998	1.001	0.984	0.991	0.986	1.002	0.994	0.976
L_{1WM}/L_{1Med}	0.943	0.925	0.873	0.924	0.852	0.888	0.901	0.941	0.940	0.957	0.925	0.936	0.931	0.887

Ami a hagyományos medián módszerrel való összehasonlítást illeti, a WM minden esetben jobban teljesített, átlagosan 10,7%-kal, maximum 20,2%-kal.

2.9 Következtetések, összegzés

Az előzőekben ismertetett hisztogramalapú súlyozott medián eljárás hatékonyságát a digitális domborzati adatok zajcsökkentésére bizonyítottam. Az eljárás fő célja az adatmátrixokban lévő kiugró zajok kiküszöbölése, különösen akkor, ha az adatpontok nagy százalékát szennyezi kiugró zaj.

A vizsgált különböző zajamplitúdók és zajterhelési százalékok átlagában a WM módszer 14-23%-kal jobb teljesítményt nyújtott a standard medián szűrési eljárásnál a különböző adatrendszereken L_1 normával mért adattávolság tekintetében a nem nulla átlagú zajok kiküszöbölésére. A módszer nulla átlagú zajok szűrésére szolgáló változata átlagosan 14,3%-kal jobban teljesített a hagyományos mediánszűrőhöz képest.

A módszer továbbfejlesztését tekintve az eljárás általános finomításán és optimalizálásán túlmenően különösen az alacsony zajexpozíciójú esetek hatékonyabb kezelése érdekében történő módosításra nyílik mód.

2.10 Tézis

3. Tézis: Kifejlesztettem egy új, hisztogramalapú zajszűrési eljárást, amely a leggyakoribb érték módszerét (Most Frequent Value-t, röviden MFV-t) felhasználja topográfiai adatok szűrésére. A fenti módszer részben az adatok dinamikus mértékű, hisztogramalapú előszűrésével, MFV-t, illetve további súlyozásokat felhasználva a natív MFV-vel és a hagyományos mediánszűréssel összemérhető eredményt produkál. Az eljárás heterogén adatrendszereken szignifikánsan jobb zajcsökkenést produkál a hagyományos mediánszűréshez képest. Megállapítottam, hogy a hisztogramalapú szűrési eljárásban kiszámított részeredmények abszolút különbsége nagy megbízhatósággal képes előre jelezni, hogy a hisztogramalapú előszűrés jó eredménnyel alkalmazható-e az adott adatablakban vagy ellenkezőleg. Ezen megállapítást az L_1-norma alapján számított hiba és a részeredmények különbségének számértéke közötti 0,99 értékű korrelációs együttható támasztja alá. A megfigyelést a szűrési eljárás eredményének súlyozása során felhasználtam módszerem hatékonyságának és robusztusságának növelésére, az adott ablakbeli MFV és hisztogramalapú szűrési eredmény értéke közötti súlyozás formájában.

Kapcsolódó saját publikáció: [S4]

3. Súlyozott mediánszűrésen és leggyakoribb érték módszerén alapuló klaszterelemzési eljárás

3.1 Bevezetés

A mérési adatok klaszterelemzése a földtudományok különböző területein kiemelkedő fontossággal bír, elsősorban a kőzetek azonosítása és tagolása területén. Ennek megfelelően nagyméretű adatrendszereken számos ilyen irányú megoldás született a korábbakban és a közelmúltban is, amely hatékonyan segítette a nyersanyagkutatást. Egy korai tanulmány [23] a klaszterelemzést ígéretes eszközként említette például a szénhidrogén-tároló zónák felismerésére. Napjainkban a klaszterelemzést a nem felügyelt gépi tanulás és a nagy adatelemzés matematikai technikáinak egyik csoportjaként említik [36], amelyet széles körben alkalmaznak a szénhidrogén-kutatás kőzettipizálási problémáiban [7], [56], [61]. Zhang és munkatársai egy nem felügyelt tanuláson és földtani változások színkeveréses kiemelésén alapuló tároló-azonosítási módszert fejlesztettek ki, ahol többféle szeizmikus attribútumot klaszterelemzéssel dolgoztak fel annak érdekében, hogy kiemeljék az olaj- és gáztárolókat mutató anomáliákat [65]. Mások nem-hierarchikus klaszterelemzést használtak az áteresztőképesség előrejelzésének segítésére dolomit- és homokkőintervallumokban az Ogallah-mezőn, az USA-ban [57], a fácies meghatározására egy nyugat-afrikai homokkőformációban lévő kút esetében a formáció áteresztőképességének előrejelzése előtt [2], valamint heterogén karbonátos tározók azonosítására egy dél-iraki olajmezőn [3]. A mélyfúrási geofizikai szelvények értelmezésének újabb alkalmazásai közé tartozik az elektrofáciesek hatékony azonosítása és litológiai osztályozása [43], [62], az ún. inverziós eljárások alkalmazásához szükséges automatikus rétegvastagság-meghatározás és a szénhidrogén-tároló képződmények szonda-válaszfüggvény konstansainak becslése [50], a heterogén (és hiányos) maglaboratóriumi adatrendszerek klaszterelemzése [51], a nem-hagyományos tárolókban a magas gázpotenciálú intervallumok azonosítása és lehatárolása [28], [49], [52].

Kurrens terület a mélyfúrási geofizikai inverzió is, ahol a különböző kőzetfizikai paramétereket közvetlenül a fúrólyuk-szelvényekből (a mélység függvényében mért in situ adatokból) nyerik ki, ezzel hatékonyan csökkentik a költséges laboratóriumi mérések számát. A klaszterelemzés megközelítés a jövőben a nem-hagyományos szénhidrogén- és geotermikus tározókra is alkalmazható, ahol a klaszterelemzéssel végzett litológiai osztályozás és a formációhatárok felismerése jelentősen segíthet az inverziós becslések pontosságának és megbízhatóságának

növelésében, így az általam kidolgozott módszernek számos potenciális egyéb alkalmazási terület mellett a földtudományi adatelemzéseket illetően is fontos alkalmazási lehetősége mutatkozik az általam vizsgált adatrendszereken túl is.

A k-középpontú klaszterelemzési módszer az adattérbeli objektumok (adatvektorok) osztályozását úgy végzi, hogy az adatmegfigyeléseket hasonlóságuk alapján előre meghatározott számú klaszterbe csoportosítja [22]. A klaszterközéppontot a klaszterelemek átlagértékeként számítja ki, ami viszonylag zajérzékeny folyamat, és csak Gauss-eloszlású adatok esetén ad optimális megoldást. Számos kísérlet történt a klaszterelemzés standard algoritmusainak módosítására, hogy robusztusabb megoldást adjanak. A gyakorlatban ez akkor lehet indokolt, ha a Gauss-statisztikára vonatkozó feltételezések nem teljesülnek, vagy a mérési adatsorok bizonyos számú kiugró megfigyelést tartalmaznak. Például az evolúciós számítást, mint globális optimalizálási technikát alkalmazzák a varianciaarány-kritériumnak nevezett alkalmassági függvény optimalizálására a klaszterek optimális belső kohéziójának és külső elszigeteltségének elérése érdekében [9]. A kiugró értékek eliminálása fontos lépésnek bizonyul a klaszterelemzés előtt vagy közben. Egyesek erre a célra a maximum likelihood módszert alkalmazták [17], míg Filzmoser és társai adaptív outlier detektálási technikát mutattak be a feltáró geokémiai megfigyelésekhez [15]. Vizsgálataim során én is a kiugró értékek eliminálásának irányába indultam el, támaszkodva a hisztogramalapú zajszűrési eljárásomban tapasztaltakra.

A k-középpontú eljárás másik – azonban az outlierektől sem független – fent említett nagy problémáját, miszerint a klaszterelemzés eredménye nagymértékben függ a kezdeti centroidok kiválasztásától, a Steiner-féle MFV eljárás bizonyos mértékig önmagában is kezeli robusztussága okán. Mélyfúrási geofizikai adatok klaszterelemzése során a két módszert vizsgálva azonban olyan megfigyelést tettem, hogy bizonyos esetekben – főként, de nem kizárólag a kezdeti klaszterközéppontok zajos volta esetében – utóbbi is nem megfelelő eredményt produkál. Ezen megfigyelés adta annak az alapját, hogy a középpontok megválasztása során a súlyozott medián eljárásnál alkalmazott szűrési módszeremet beillesszem a folyamatba.

A Steiner-féle leggyakoribb érték (MFV) eljárás klaszterezésre történő alkalmazásának jellemzője, hogy az adott centroidot a klaszterelemek MFV-jeként (leggyakoribb értékeként) számolja ki (a k-means eljárásban alkalmazott átlag helyett), és a klaszterelemek középponttól való teljes eltérését súlyozott euklideszi (Steiner) távolsággal határozza meg. Mindezt k számú centroidra hajtjuk végre. Nem csupán a klaszterközpontok, hanem a klaszterek közötti

távolságok is súlyozottak a kiugró adatok káros hatásainak csökkentésére. A folyamat a klaszterelemek teljesen automatikus súlyozására támaszkodik, amely nem igényel megkötést a megfigyelt változók statisztikai eloszlására vonatkozóan.

3.2 A Steiner-féle MFV eljáráson alapuló klaszteranalízis

Az általam főként a szénhidrogén kutatás céljára megalkotott robusztus klaszterezési eljárásban a következőkben ismertetésre kerülő alapeljárást fejlesztettem tovább. A módszert megalkotó vizsgálatokat földtudományi mélyfúrási adatokon kezdtem, így ezek segítségével ismertetem a kiindulópontot jelentő, klaszterezésre alkalmazott MFV eljárást.

Általános jellemző a mélyfúrási geofizikai adatokat tekintve, hogy az aktuális (**D**) adatmátrix d_{il} eleme az l-edik fizikai változó értékét jelzi az i-edik mélységi koordinátánál (l=1,2,..., *L*, ahol *L a* rögzített mérési szelvények száma). Ez a módszereket ismertető mind szintetikus, mind terepi mélyfúrási példákra igaz, azzal a kiegészítéssel, hogy a több fúrásból álló alkalmazásokban az *i* index nem csak egy, hanem több fúrás mélységintervallumán is átfut (i=1,2,...,*N)*. Az adatmátrix i-edik sorában a $d^{(i)}$ vektor az L-dimenziós adattérben egy objektumot képvisel, amely többféle adattípust is tartalmazhat, például nukleáris, elektromos, radioaktív, akusztikus és egyéb típusú adatokat. A $d^{(i)}$ és a $d^{(j)}$ adatobjektumok közötti hasonlóságot általában a Minkowski-távolsággal lehet mérni

$$\delta^{(M)} = \left[\sum_{l=1}^{L} \left| d_l^{(i)} - d_l^{(j)} \right|^p \right]^{1/p}, \quad (27)$$

ahol *p=1* a Manhattan-távolságot (cityblock), p=2 pedig az euklideszi távolságot jelenti. Ez utóbbi norma csak akkor ad optimális megoldást, ha az adatok Gauss-eloszlást követnek. Ha a megfigyelések nem Gauss-eloszlásúak, a Manhattan-távolság jobban teljesít, mivel kevésbé érzékeny a kiugró értékekre.

A robusztusabb megoldás érdekében a klaszterezésre módosított MFV eljárás ehelyett az adatobjektum és az adott klaszter középpontja közötti súlyozott, Steiner-távolságot számol:

$$\delta^{(\mathrm{St})} = \left[\left(\sum_{l=1}^{L} \phi_l \right)^{-1} \sum_{l=1}^{L} \phi_l \left(d_l - c^{(\mathrm{MFV})} \right)^2 \right]^{1/2}, \quad (28)$$

ahol a súlyozási együtthatókat (φ) a Steiner-módszernél ismertetett súlyfüggvény számítás alapján kell kiszámítani. (Formálisan, ha a súlyozási együtthatók egyenlőek 1-gyel, és a centroidot a klaszterelemek átlagaként számoljuk, a fenti képlet az euklideszi távolságon alapuló K-means klaszterelemzési (CA, Cluster Analysis) módszerhez vezet).

Az csoportosítási eljárás során minden egyes iterációban minden objektumot az új klasztert alkotó legközelebbi centroidhoz rendelünk. A tényleges konfiguráció iteratív módon javul a centroidok pozícióinak újraszámításával. Ha ezek nem változnak jelentősen, a klaszterezési eljárás leáll.

3.3 A módosított WMFV klaszterelemzési eljárás

Mind a fent ismertetett, klaszterezés céljára módosított MFV, mind az általam tovább módosított MMFV (Modified MFV, MMFV) klaszterezési módszer klaszterközéppontokat számol, majd az adatokat a klaszterközéppontok köré partícionálja. Különbség viszont az MFV és MMFV módszerek, illetve értékek között egyrészt, hogy az MFV esetében a kiinduló klaszterközéppont mátrix a cityblock metrikát felhasználó k-means eljárás által adott mátrix, az MMFV esetében pedig a fent bemutatott, hisztogram műveletek által az aktuális partíciókból előállított halmazok mediánjai által adott mátrix. Másrészt, míg az MFV az adatok aktuális centroidhoz való partícionálása során a partíció vektorok elemeit szorozza a módszer által kapott súlyokkal, ez a súlyozás az MMFV esetében nem történik meg.

A WMFV (Weighted MFV) módszer által adott végleges klaszterközéppontok az MFV-technika és a hisztogram-megközelítéssel segített MMFV-módszer eredményének súlyozott átlagaként adódnak megoldásomban: MFV ill. MMFV módszerek által adott klaszterközéppont értékek w_1 és w_2 súlyok által súlyozott átlagai lesznek [S7].

Az eljárásban az MMFV centroidszámítás módszere is kismértékben módosításra került a klaszterezési feladatra történő átalakítás során a korábban a szűrési módszernél ismertetettekhez képest. Az m_{e1} paraméter értékének meghatározásához első lépésben adatszűkítési folyamat történik. Ebben a folyamatban az ablakelemeket először érték szerint rendezzük, majd az értékkészlet alapján hat egyenlő szélességű tartományra (binre) osztjuk. Itt kiszámítjuk a λ_1 arányt, mint az adatablak elemszámának és a legmagasabb értékű tartomány nélküli tartományok elemszám összegének arányát. Ezután kiszámítjuk a λ_2 értéket, amelynek értéke a legnagyobb és a második legnagyobb elemszámú tartomány elemszámának aránya lesz. Ha $\lambda_1 > \lambda_2$: akkor a legmagasabb értékű (a legnagyobb értékekkel rendelkező) tartományt, ellenkező esetben a legnagyobb elemszámmal rendelkező tartományt választjuk a kiválasztott halmaznak (D). Így a nagyobb λ értéket az élesebb vágás jeleként használjuk. Végül m_{e1} ennek a D halmaznak az átlaga lesz.

Az m_{e2} értékét hasonlóan számítom ki, azonban a tényleges adatablakot nem 6, hanem 8 részre osztva. A fentiekből látható tehát, hogy λ_1 és λ_2 ez esetben másképp számítódik, mint a

módszer pusztán zajszűrésre használt esetében (λ_3-hoz ill. λ_4-hez hasonlóan), ill. tapasztalati alapon más az osztásszám. Az ezt követő részsúly számítások a korábbiakban ismertetett módon történnek, λ_3, λ_4 arányszámok, illetve **w** súlyok előállítása azonban módosul az alábbiak szerint. Az aktuális adatablakban lévő elemek értéktartományát két és három egyenlő tartományszélességű halmazra osztjuk, majd előállítjuk a két arányszámot:

- λ_3: a két tartomány közül a nagyobb és a kisebb elemszámú tartomány közötti elemszám-arány,
- λ_4: a legnagyobb és a második legnagyobb tartomány aránya a háromból (ismét az elemszám tekintetében).

Ha $\lambda_3 > \lambda_4$, akkor az új halmaz (*E*) a 2 tartomány közül a nagyobb elemszámú, egyébként a 3 tartomány legnagyobb elemszámú tartománya.

Ezután összehasonlítjuk a csonka *D* és *E* halmazok méretét. Ha a *D* mérete nagyobb, mint az *E* mérete (a több részre való felosztás ellenére), akkor ez azt jelzi, hogy az aktuális adatablak adatai nem jól strukturáltak a hisztogramalapú szétválasztáshoz, ezért a hisztogramalapú MFV eredménynek némileg kisebb súlyt kell kapnia, míg fordított esetben ennek ellenkezően. Annak érdekében, hogy ezt a súlyozást el tudjam végezni, a w_1 súlyérték az előbbi esetben 10, az utóbbi esetben pedig 1,25 értéket kapott. Ezen két érték tapasztalati alapon került beállításra. Mivel w_2 értéke jellemzően 1 környékén alakul, így w_1 10, illetve 1,25-ös értékével biztosítható, hogy a módosítás nélküli MFV érték kellő súlyt kapjon az eredményben, amikor szükséges (amikor a fentiek szerint a hisztogramalapú változat önmagában nem ideális eredményt adna).

Ezen a ponton már ismerjük az aktuális adatpont **w** súlyvektorát:

$$\mathbf{w}=[w_1 \;\; w_2]^{\mathrm{T}}. \qquad (29)$$

A **w** súlyvektorban (29. egyenlet) a w_1 és w_2 súlyoknak egyrészt az aktuális adatablak eredeti MFV értékére, másrészt a javított MFV értékre (MMFV) hatnak, w_1 az utóbbit, w_2 pedig az előbbit súlyozza az alábbiak szerint, előállítva a klaszterközéppontot is:

$$c_k=(w_1 * \mathrm{MMFV}+w_2 * \mathrm{MFV})/(w_1+w_2). \qquad (30)$$

Mind az MFV, mind a hisztogramalapú MFV módszer klaszterközéppontokat számol tehát, illetve az adatokat a klaszterközéppontok köré partícionálja, majd pedig a kettő közötti súlyozást valósítok meg az MFV és MMFV súlyozásával előálló WMFV (Weighted MFV) eljárásban.

3.4 A WMFV klaszterelemzés alkalmazása

A bemutatott WMFV klaszterelemzés módszert először egyetlen fúrás adataira alkalmaztam, majd az eljárást kiterjesztettem több fúrás adataira, hogy rekonstruálható legyen a klaszterek többdimenziós térbeli eloszlása, amely feltárja a vizsgált képződmények litológiai és petrofizikai jellemzőit. Elsőként egy szintetikus, majd egy nagy in situ adatrendszer került elemzésre, amelyet egy hazai szénhidrogén mező fúrásaiban mértek (Szabó et al. 2023). A klaszterelemzés célja a fúrással harántolt összetett miocén korú szénhidrogén-tároló összlet vizsgálata, a gáz- és víztároló zónák elkülönítése. A terepi eredmények pontosságát mag áteresztőképesség-mérések és független fúrólyuk szelvényértelmezések igazolták, amelyek mindezeken túl egybeestek egy általam megalkotott, a klaszterezési módszer(ek) zajszűrő képességét jellemző gradiens metrika által adott eredményekkel [S7]. A fentiek szerint tehát az elsődleges vizsgálatok mélyfúrási adatokon történtek, sorrendben szintetikus, majd valós (terepi) 1 dimenziós, ill. terepi 2 dimenziós adatok segítségével.

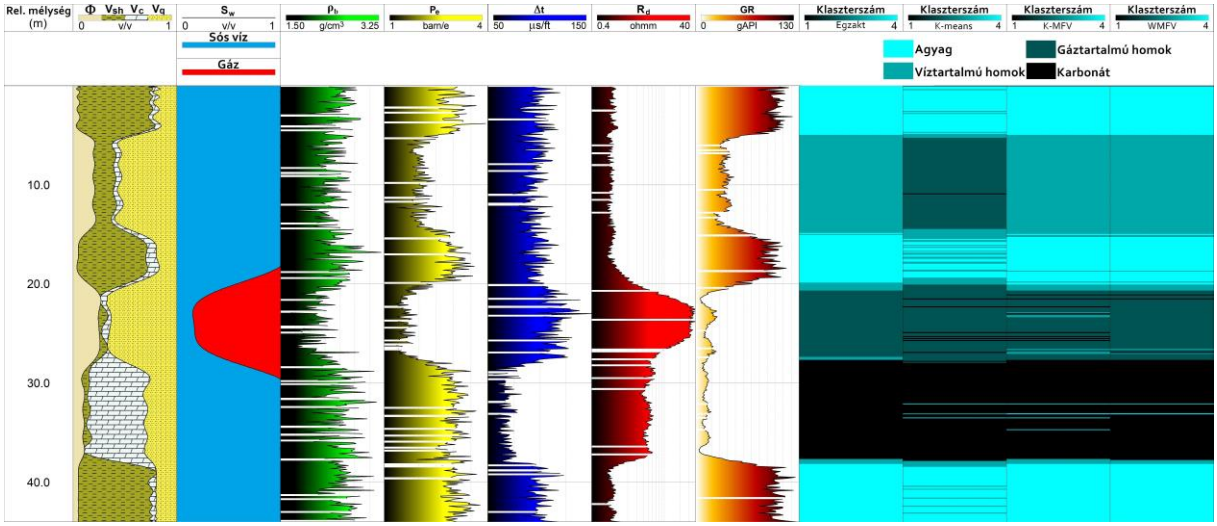
3.4.1 Szintetikus mélyfúrási adatrendszeren történő vizsgálatok

A WMFV módszer hatékonyságának tesztelésére elsőként egy szintetikus modellezési kísérletet végeztem. Ehhez egy fiktív (pontosan ismert) petrofizikai modellt építettem. A valós geofizikai mérések szimulálása érdekében ezen szintetikus (zajmentes) adatokhoz tetszőleges mértékű zajt adhatunk tesztelve, hogy mennyire pontosan és megbízhatóan rekonstruálja a WMFV módszer a klaszterek (ismert) mélységi eloszlását.

Első lépésben a fent említett kiinduló rendszerként egy hatrétegű inhomogén modell alapján zajmentes adatrendszert számítottam. Ez a következő számított adatértéket tartalmazta: a kőzetformáció sűrűsége (ρ_b), a természetes gamma-sugárzás intenzitása (GR), a fotoelektromos abszorpciós index (P_e), a szónikus P-hullám terjedési ideje (Δt) és a látszólagos fajlagos ellenállás (R_d). Az adat-modell kapcsolatokat leíró szondaválaszfüggvények és konstansai a földtudományban szokásos és ismert formában kerültek megalkotásra [S7], melyek az értekezés szempontjából nem hordoznak lényegi információt, így nem ismertetem (az [S7] publikációban azonban megtekinthetőek).

A fenti mélyfúrási geofizikai adatrendszerhez ezt követően 10 %-os Gauss-eloszlású zaj került hozzáadásra. Ezen kívül véletlenszerűen kiválasztásra került az összes adat 1/20 része, és törlésre került (11. ábra 3-7. sávjainak fehér csíkjai), majd multilineáris regressziós imputációs módszerrel pótlásra került a Szabó és munkatársai által megalkotott módszer szerint [51] (ezzel további zajt is eredményezve egyben az adatmátrixon). Az egzakt, ismert petrofizikai modellt

a 11. ábra első két sávja tartalmazza, míg a klaszterelemzés bemenő adatait (10%-os Gauss eloszlású zajjal szennyezett, nem teljes bemeneti adatsor) a 3-7 sávok mutatják.



11. ábra. A szintetikus adatok klaszterelemzésének eredménye, 1-2. sáv: petrofizikai modell, a 3-7. sávok: bemeneti adatsor, a 8. sáv: klaszterek pontos helye, a 9-11. sáv: k-means, MFV, ill. WMFV eljárásokkal becsült eredmények.

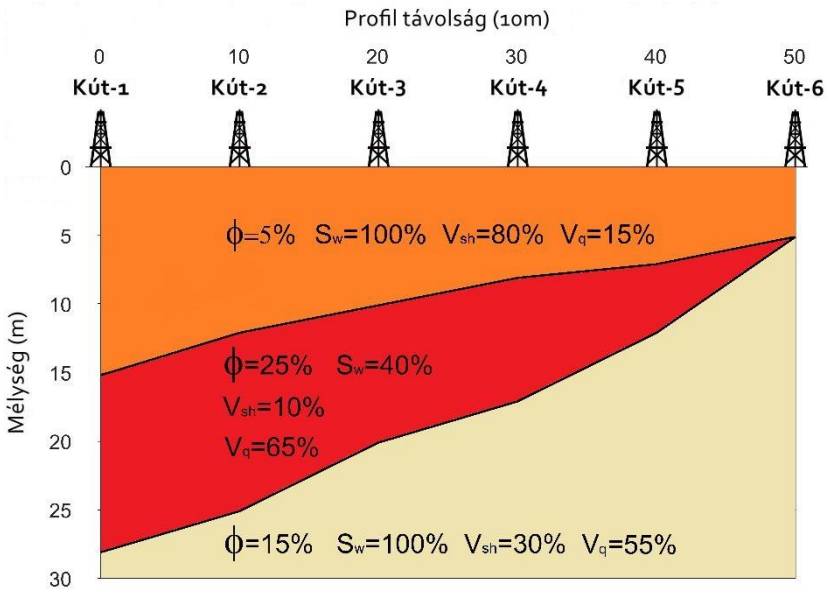
A klaszterelemzés eredményét a 8-11. sávok tartalmazzák a 11. ábrán, amelyek mindegyike esetében könyök-módszer által és szakértői tudás alapon is 4-es input klaszterszám volt a feltételezett. A zajmentes adatok klaszterelemzése a 8. sávon a négy klaszter függőleges eloszlását (azaz a klaszterszámok mélység szerinti változását) adja, ami megegyezik a pontos megoldással, a kőzettípusok ugyanis a feldolgozott intervallum mentén a kis mélységtől kezdve a következők: 5 m vastag agyag, 10 m vastag víztartó homokkő, 5 m vastag agyag, 8 m vastag szénhidrogén-tároló homokkő, 10 m vastag mészkő és 6 m vastag agyag. A 11. ábra további sávjait tekintve, a 9-11. sorrendben a k-means, MFV, WMFV eredményeket jeleníti meg. Az adatgyűjtés helyeit a relatív mélységskála méterben mutatja.

A mészkőréteghez fekete színt társítottam, a meszes üledékeket világosabb szín jelzi, míg a szénhidrogén-telített homokot sötétkékkel ábrázoltam. A k-means eljárás mellett a normál MFV (más néven K-MFV a klaszterezésre való alkalmazás okán), és módszerem, a módosított MFV klaszterező eljárás (WMFV) is lefuttatásra került. Ezek eredményeit is összehasonlítva látható, hogy a hagyományos k-means módszer kifejezetten érzékeny az adatzajokra, sok esetben összekeveri a víztartalékot a szénhidrogéntartalékkal, a réteghatárok mélyebben jelentkeznek, és az effektív rétegvastagságra vonatkozó becslése nem pontos. Ez utóbbi szembetűnő a második és negyedik rétegek esetében. Megfigyelhető az is, hogy a módosított

MFV eljárás kimenete a legkevésbé zajjal terhelt, és ez határozza meg legjobban a tározó zónákat.

3.4.2 2D szintetikus mélyfúrási adatokon történő vizsgálatok

A vizsgálatok folytatásaként a robusztus klaszterezési módszert többdimenziós alkalmazásokra terjesztettem ki, több, a földfelszínen egy nyomvonal mentén fúrt kút összes adatának együttes feldolgozásával. Ehhez egy 2D-s háromrétegű kőzetfizikai modell került felépítésre, amely a 12. ábrán látható, ahol Φ porozitást, S_w víztelítettséget, V_{sh} agyagtartalmat, V_q-homok részarányt jelöl. Fentről kezdve a szekvencia egy agyag, egy szénhidrogén-tároló homok és egy kisebb hézagtérfogatú (az előzőnél nagyobb agyagtartalmú) víztároló homokformációt foglal magában. A fentiek előállítási képleteinek ismertetését a dolgozat az előbbihez hasonlóan nem tartalmazza.



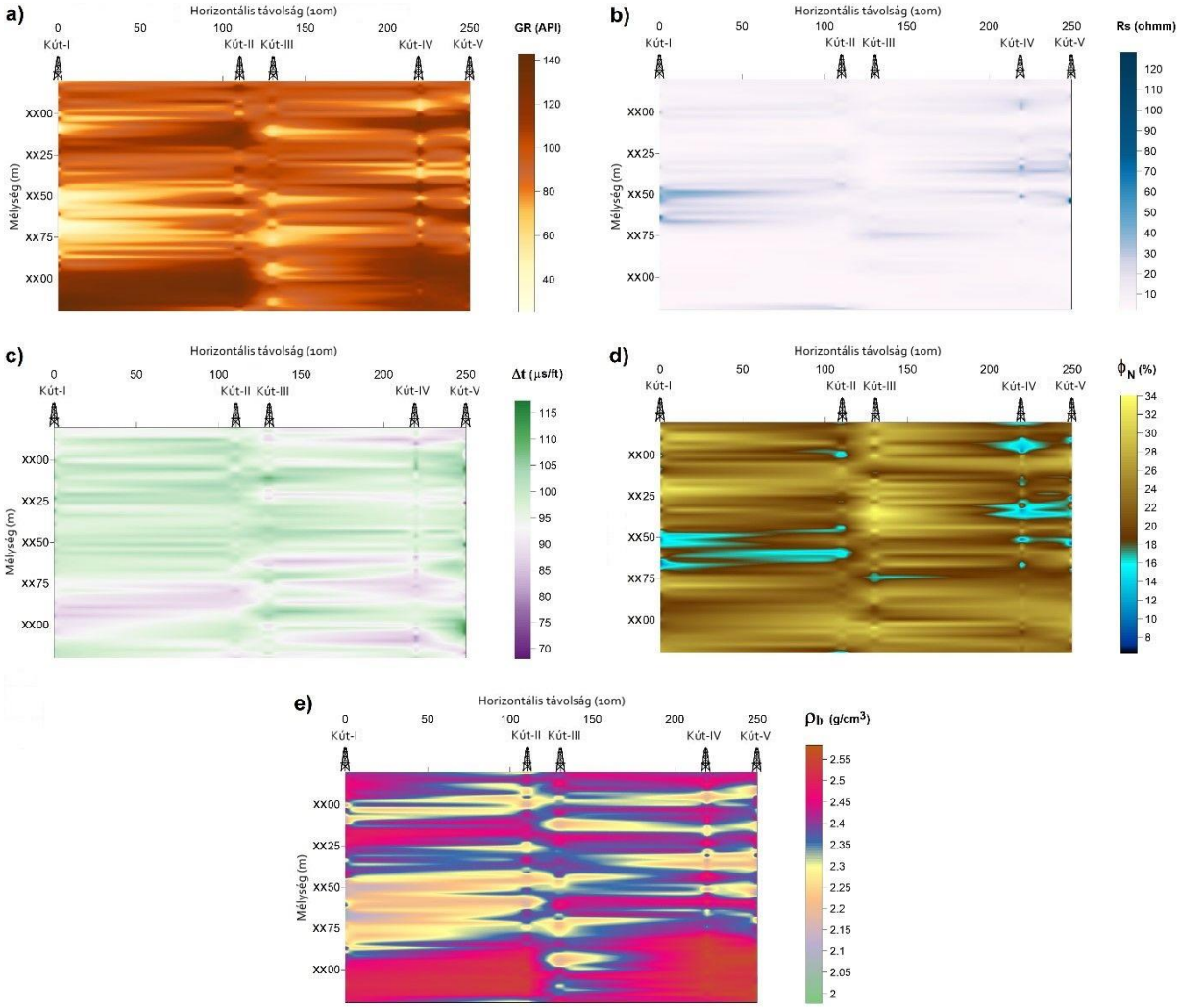
12. ábra Modellezett, szintetikus adatrendszer felépítése

A függőleges irányú mintavételi távolságot 0,1 m-re, a fúrólyukak (kút-1-kút-6) közötti 100 m-es távolságot választva összesen 8970 adatpont állt elő a vizsgálatokhoz. Az 5 %-os Gauss-zajjal szennyezett mélyfúrási geofizikai adatrendszert a 12. ábra 2D szelvények formájában szemlélteti.

3.4.3 Terepi adatok 2D klaszterelemzése

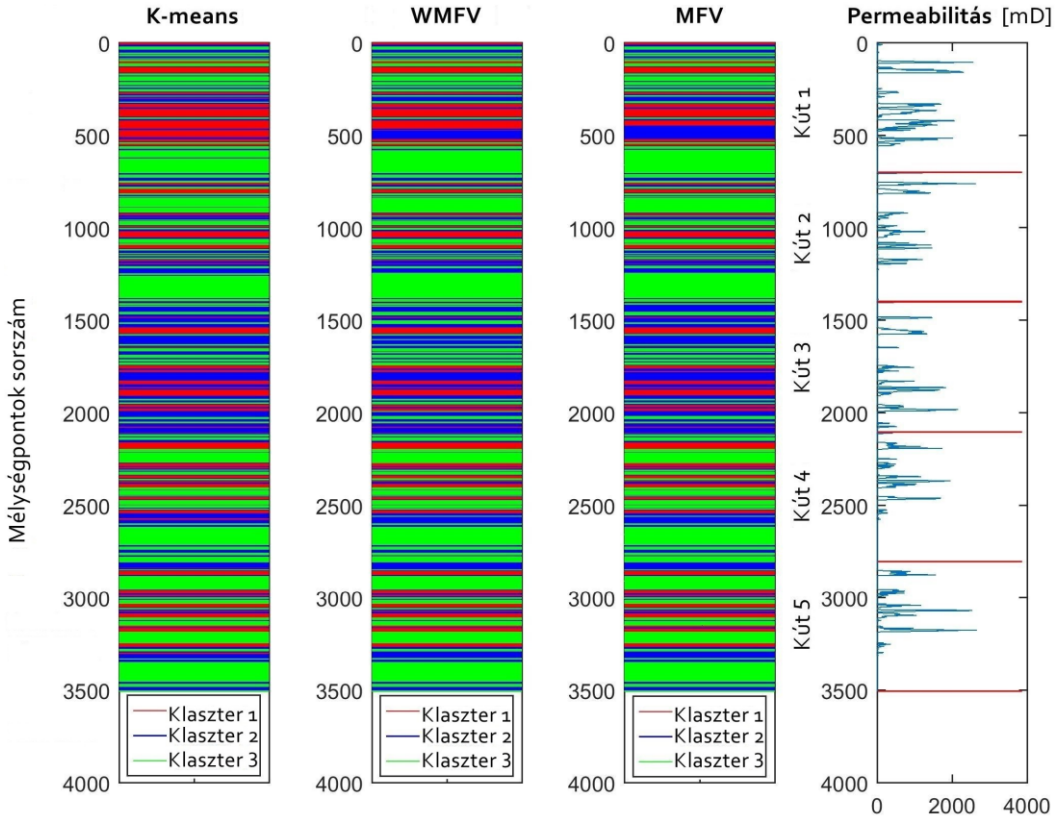
Módszeremet egy kelet-magyarországi gázmezőn gyűjtött terepi adatok segítségével is vizsgáltam. Az adatrendszer egy 140 m hosszú, miocén agyagos (iszapos) homokos szekvenciák mélységintervalluma mentén egyidejűleg öt fúrásból (kút-I-kút-V) mért fúrólyukszelvény készlet volt. A fúrások vízszintes koordinátái egy vonal mentén a következők

voltak: 0 m, 1100 m, 1300 m, 2200 m, 2500 m. A bemeneti változók a természetes gamma-sugárzás intenzitása (*GR)*, a sekély behatolású szondával mért fajlagos ellenállás (R_s), az akusztikus terjedési idő (Δt), a neutron-porozitás (ϕ_N) és kőzetsűrűség (ρ_b) voltak, melyek a 13. ábrán láthatóak.



13. ábra. A mért fúrólyukszelvények interpolált szelvényei, a) természetes gamma-sugárzás intenzitása (GR), kisepert zóna fajlagos ellenállás (R_s), c) akusztikus terjedési idő (Δ t), d) neutron-porozitás (ϕ_N), e) kőzetsűrűség (ρ).

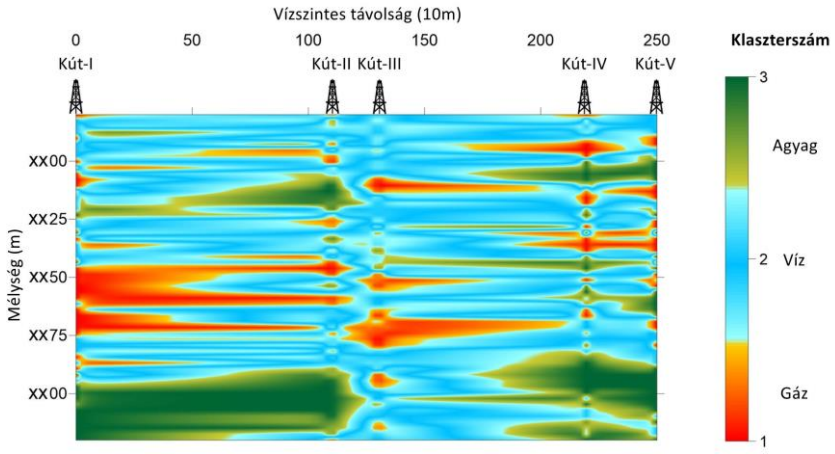
A keresett lehetséges tárolókat az adatok elemzésével szakértői tudást felhasználva a 13.b ábrán sötétkék színnel jelölték földtudományi szakemberek. Három klasztert feltételeztek: gáztározókat (1. klaszter), vízzel telített homokokat (2. klaszter) és agyagrétegeket (3. klaszter). Az általam kifejlesztett klaszterezési módszerek eredményeit a 14. ábra szemlélteti.



14. ábra. A k-means (első sáv), a MFV (harmadik sáv) és a továbbfejlesztett MFV (WMFV) módszerek (második sáv) által becsült klaszterek egydimenziós ábrázolása, illetve a független értékelésből származó áteresztőképesség (utolsó sáv), amely jó korrelációt mutat a klaszterelemzés eredményeivel.

Vizuálisan vizsgálva a k-means, a tiszta MFV-alapú (MFV) és a javított MFV-alapú (WMFV) klaszterelemzés ez esetben csaknem azonos eredményeket ad. Láthatók a gáztároló zónák a magas permeabilitás értékeknél (1. klaszter). A kék intervallumok (2. klaszter) a közepes/magas permeabilitású, sós vízzel telített képződményekhez társíthatók, míg a zöld intervallumok (3. klaszter) a vízáteresztő képességgel nem rendelkező agyagos képződményeket mutatják. Ettől lényegibb információ azonban azon szakértői megállapítás, hogy a litológiai hatás csökkentésével jó korreláció adódik az így kapott klaszterezési eredmények és az abszolút permeabilitás nagysága között, így ez a módszerek összehasonlításának egy jó összehasonlító mérőszámaként adódik [S7].

Az azonosított klaszterek és a permeabilitás közötti távolsági korrelációs értékek pedig -0,62 (k-means), -0,63 (MFV), -0,64 (javított MFV, vagyis WMFV) voltak, tehát ezen összehasonlításban eljárásom adta a legjobb eredményt. A 15. ábrán a WMFW eljárás által kapott klaszterezési eredmények tekinthetők meg, az ismertetett rétegek elhelyezkedését megmutató formában.



15.ábra. WMFW eredmények a különböző régetek elhatárolásának céljával

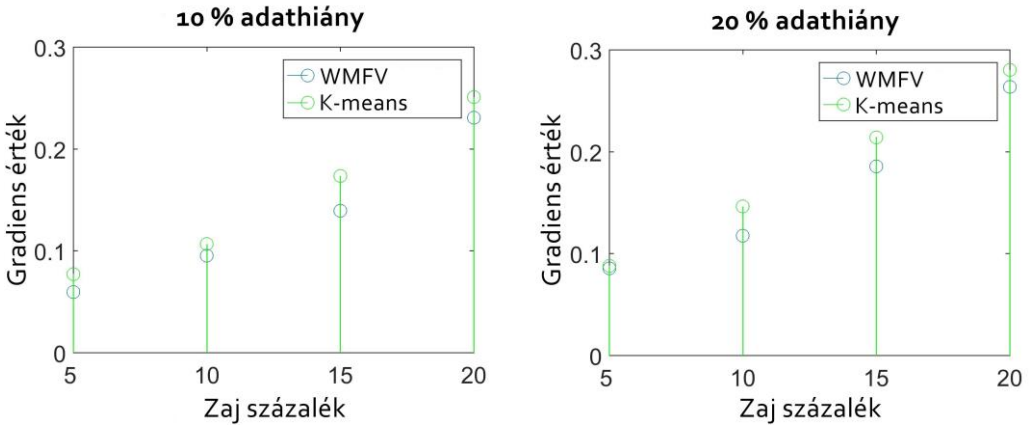
3.4.4 Validálás ismételt futtatások alkalmazásával

A fentieken túl azt is vizsgáltam, hogy a klaszterezési módszerek összehasonlítási eredményei milyen mértékben konzisztensek ugyanazon adatrendszerre történő több futtatás, illetve különböző mértékű zaj, és adathiány mértékek esetén. Ehhez a 11. ábrán már ismertetett 1D adatmodellhez nyúltam vissza. Különböző bemeneti adatkészleteket hoztam létre úgy, hogy a zajmentes adatokhoz 5%, 10%, 15% és 20% Gauss-eloszlású zajt adtam, illetve ugyanazon zajszint mellett 1/20 és 1/10 adatvesztést is figyelembe vettem.

A kimeneti klaszterindex vektorok változékonyságának mérési metrikájaként gradienst használtam a következő formában:

$$grad(v) = \frac{\sqrt{\sum_{i=1}^{N}\left(\frac{dv}{dz}\right)_i^2}}{N}, \quad (31)$$

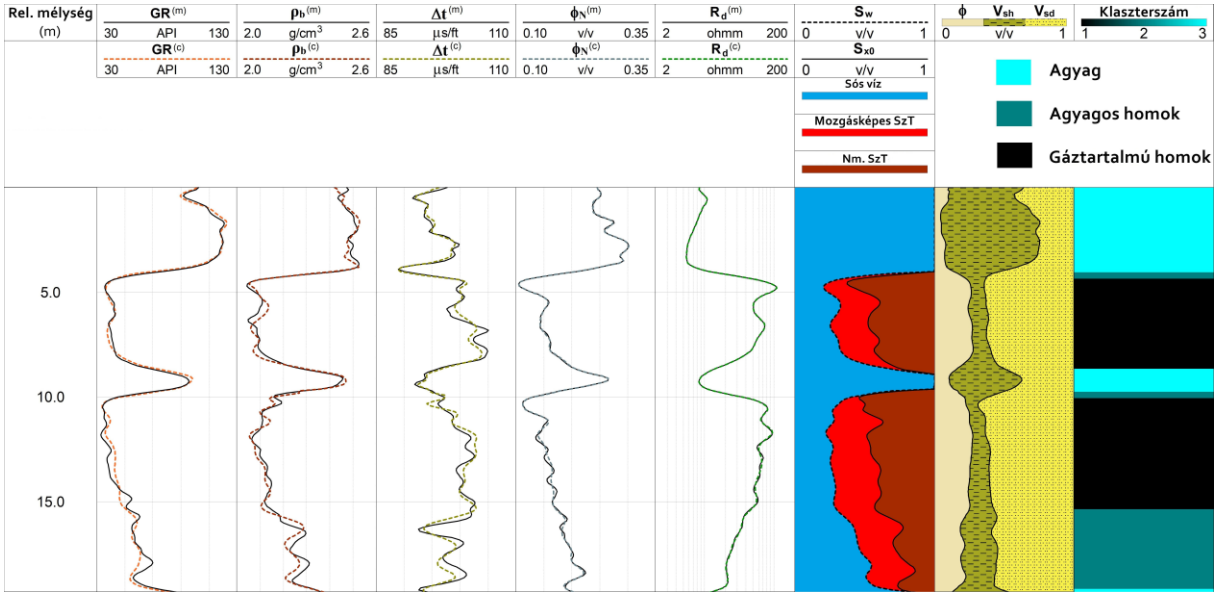
ahol v_i a **v** klaszterindex vektor (i=1,2,...,*N*) i-edik eleme. A klaszterindex vektorok gradiens értékeit a 15. ábrán ábrázoltam, azok változékonyságát bemutatandó. Az ábrán az egyes értékek rendre 10 számítógépes futtatás eredményeinek átlagát jelentik. A WMFV módszer variabilitása kék színnel, a k-means CA eljárásé pedig zöld színnel látható. Az eredmény azt mutatja, hogy a WMFV eljárást kevésbé befolyásolják az adatzajok, ami megerősíti a 11. ábrán látható eredményeket. Az eredmény feltételezhetően független az egzakt modell klaszterszám-eloszlásától, a gradiens értéke a klaszterelemzés eredményének bizonytalanságát fejezi ki (a 11. ábrán a klaszterindex vektorok hibás vékony csíkjainak sűrűsége), amelyet a bemeneti adatok különböző pontossága mellett kapunk.



16. ábra. A klaszterezési módszerek ellenőrzése a k-means és a WMFV eredményekből (ordináta tengely) származó gradiens értékével különböző Gauss eloszlású zajszintek (abszcissza tengely) a bemeneti adatmátrix adathiány különböző mértékei mellett.

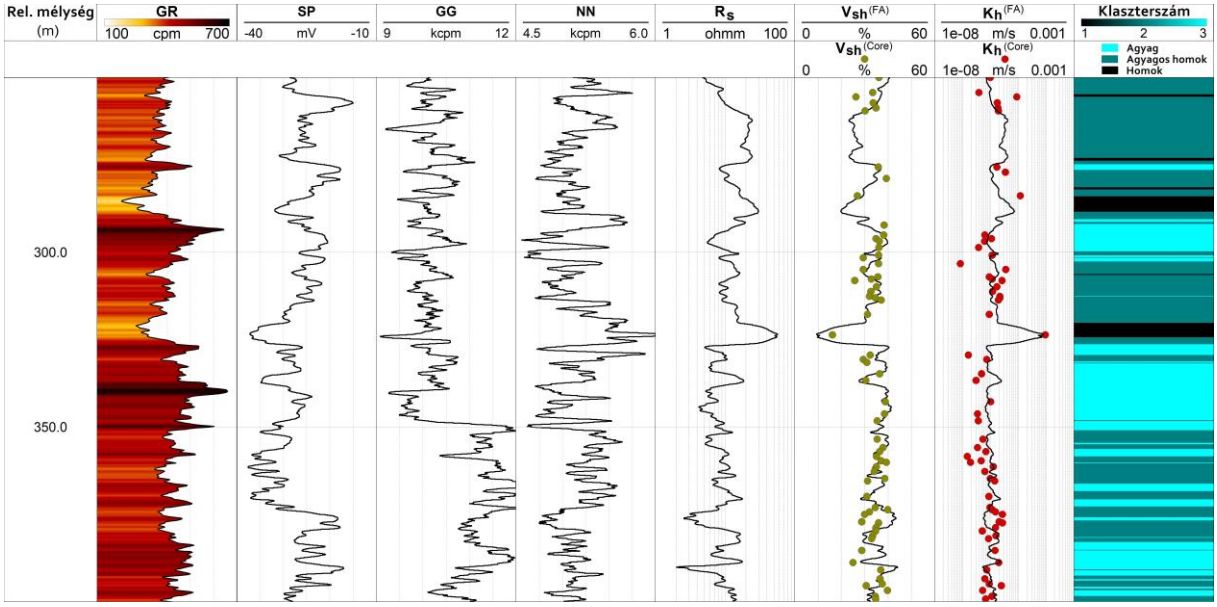
3.4.5 További vizsgálatok

Klaszterezési módszerem további két, független kiértékelt adatrendszeren is validáltam. A vizsgált magyar szénhidrogénmezőn az WMFV megközelítéssel kapott eredmények közvetlenül összehasonlíthatók a szakértők által ismert magadatokkal, vagy a kútfúrások kvantitatív értelmezéséből nyert petrofizikai paraméterekkel [S7]. Ezeket egyetlen kutatófúrás egy 18 méteres szakaszán mért, ill. ún. inverzióval meghatározott paramétereket a 17. ábra 1-5 sávjai mutatják be.



17. ábra. A klaszterelemzés eredménye egy kútban az 1-5. sávok a mért (m) és az inverzióból származtatott számított adatok, a 6. sáv a víz és a szénhidrogén telítettségét, a 7. sáv a porozitás, az agyagtartalom és a homok térfogatát, a 10. sáv a WMFV eljárással kapott klaszterezési eredményt mutatja.

Hasonló egy kutas, ám az előbbitől eltérően termálvizes (kelet-magyarországi) kút esetében tartalmazó adatrendszerre történő futtatás esetét ismerteti a 18. ábra. A kőzetminta adatok szakértői értelmezése, illetve a klaszterezés kimenete által együttes kiértékeléssel azonosítani lehet a jó vezetőképességű zónákat (1. klaszter), és el lehet különíteni őket az alacsony áteresztőképességű (2. klaszter) és a vízzáró rétegektől (3. klaszter), így ilyen jellegű adatok esetében is használhatónak mutatkozott a módszer.



18. ábra. A klaszteranalízis eredménye egy magyarországi vízkútban, az 1-5. sávok a megfigyelt értékeket mutatják, a 6. sáv tartalmazza a faktorelemzéssel és magmérésekkel becsült palatérfogatot, a 7. sáv a faktoranalízissel és magmérésekkel előre jelzett vezetőképességet, a 8. sáv a WMFV klaszterezés eredményét mutatja.

3.4.6 Tézis

4. Tézis: Kifejlesztettem egy leggyakoribb érték (MFV) - alapú klaszterelemzési módszert, ahol a kezdeti centroidokat hisztogram-alapú szűréssel határoztam meg. A módszert mélyfúrási geofizikai adatokra alkalmaztam és meghatároztam a klaszterek mélységi változását. A kidolgozott súlyozási eljárással hatékonyan megtisztítottam az eredményt a kieső értékektől. Ezen felül bemutattam, hogy a módszer eltérő zajszennyezettség és adathiány mértékek esetében is robusztus, megfelelő minőségi jellemzőkkel bíró csoportosítási alternatívát jelent. Az eljárást több fúrás adatrendszereinek együttes feldolgozására is alkalmassá tettem. A robusztus eljárás javítja a klaszterelemzés zajérzékenységét, biztosítja a kőzetrétegek megfelelő felbontását, segíti a nyersanyagok azonosításának hatékonyságát és a mérési adatok földtani-geofizikai értelmezését.

Kapcsolódó saját publikáció: [S7]

3.4.7 A WMFV eljárás alkalmazása, validációja hidrológiai adatokon, Hurst exponensen alapuló mérőszám bevezetése

A korábbiakban ismertetett mélyfúrási adatokon túl egy ezektől eltérő szerkezetű adatrendszeren is validáltam klaszterezési módszerem [S8]. A Nílus folyó vízszint ingadozásai képezték a vizsgálataim tárgyát. Ez az ember által valaha mért leghosszabb idősor. Az összehasonlítások eszközkészletét tágítandó a validációkban a korábbiakon túl egy új, a Hurst-indexen alapuló mérőszámot is bevezettem mindezekhez.

A validáció céljául szolgáló adatrendszer a Nílus folyó vízszintjéről i.sz. 622–1922 között gyűjtött éves rekordok formájában állt rendelkezésre, amely egyedülálló nagyméretű adathalmazt jelentett a módszerek további összehasonlításához. Az adatsor értékeit az éves vízszintek maximum és minimum értékei [59] adták.

A klaszterezési eljárások eredményeinek vizsgálatára használt három mérőszám közül az első metrika a klaszterezési eljárásokban bevett sziluett módszer volt [39]. Második mérőszámként gradiens értéket használtam (amely gyakorlatilag a klaszterindex vektorának változékonyságát méri). Ezt a súlyozott medián módszernél, illetve a klaszterezés előbb ismertetett példáiban alkalmazottak szerint számítottam, azzal a különbséggel, hogy ez esetben nem a gradiens, hanem az (1-gradiens) értéket használtam fel. Harmadik mérőszámként a Hurst-exponenst alkalmaztam [27]. Ennek számításának első lépése a kumulált adatok átlag körüli ingadozásainak terjedelmének meghatározása (amit R-rel jelölünk), majd ennek standardizálása az adatok S szórásával elosztva. Ennek szokásos jelölése RS, ami tehát az R/S hányados.

A Hurst-exponens definíció szerint ezekből az alábbi módon számíthatjuk:

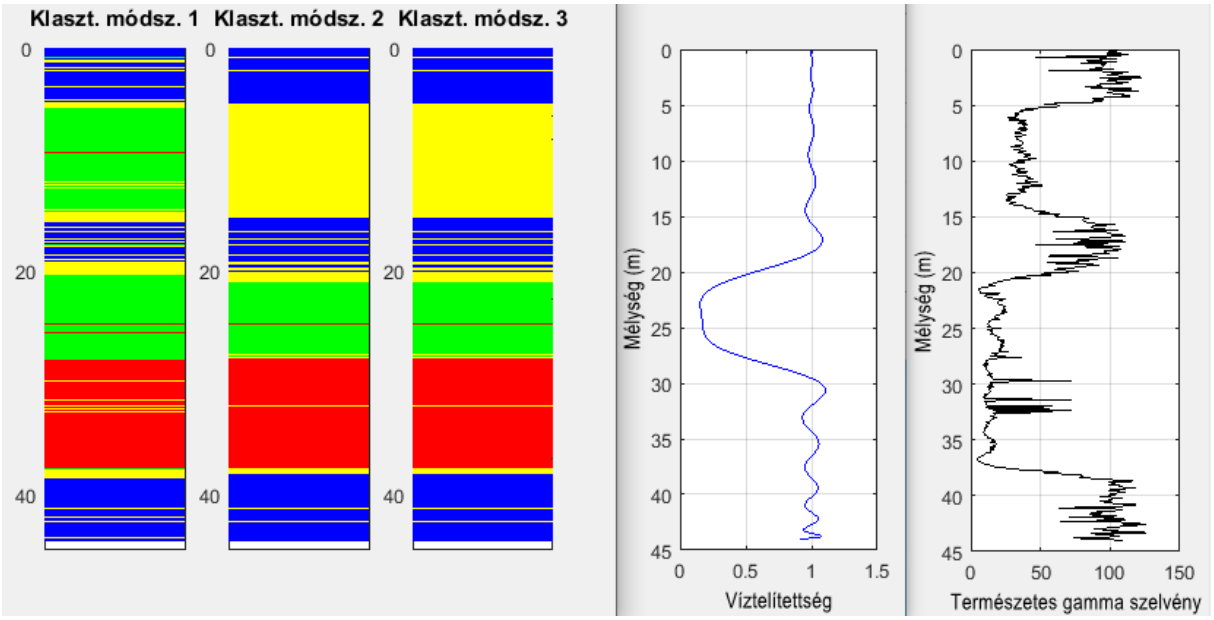
$$H = \frac{\log(RS)}{\log(N)}, \quad (32)$$

ahol N a sorozat hossza. A H értéke tehát a log-log plot meredekségét mutatja, értéke [0, 1] intervallumon változhat. Az exponens alkalmazására láthatunk egy példát Szűcs és szerzőtársai cikkében [55].

A mérőszámnak azon tulajdonságát használtam fel, miszerint amennyiben a Hurst-exponens nagyobb értékű egy adott adatvektoron egy ugyanolyan méretű másik vektorhoz képest, az esetben a vektorban nagyobb mértékben találhatók (legalább időszakos) trendek (nem pedig egy adott érték körül oszcillál). Azaz, a Hurst-exponens nagyobb értékével az adott klaszterindex vektor trendnélküli oszcillálásának kisebb mértékét mutattam ki. Ezzel együtt, az értéket nem önmagában használtam fel, hanem az adott klaszterindex vektorára kapott Hurst-

exponens érték/az eredeti adatsor vektorok Hurst-exponens értékeinek maximuma hányadosaként.

A bemutatott, Hurst-exponenst tartalmazó mérőszám adta eredmények, illetve a mérőszámok jóságának vizsgálatára először mélyfúrási geofizikai adatokat tartalmazó szintetikus adatsort használtam, különböző zajszintek mellett.



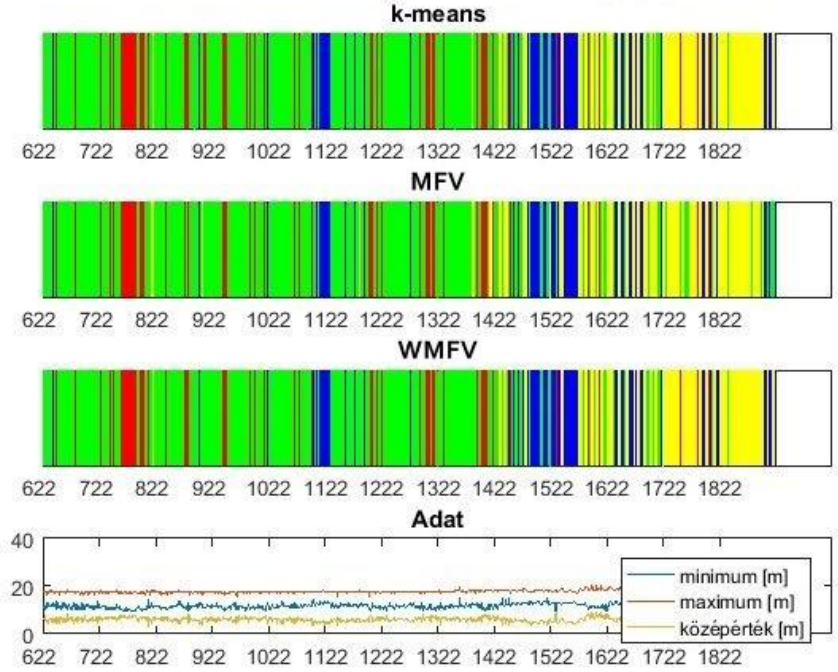
19. ábra Klaszterezés eredmények mérőszámok összhehasonlító vizsgálatához

A 19. ábrán látható, hogy a második és harmadik (a példában egymáshoz képest egyforma eredményt produkáló) klaszterezés produkálta a legjobb eredményt az adatsoron (a 20-25 méteres szakaszon található szénhidrogén-tároló réteg a példában). Ez egybeesik azzal, hogy a számszerű eredményeket mutató – a minősítő számadatok validációjára szolgáló – 20. táblázatban mindhárom mérőszám (első sor: sziluett átlag, második sor: sziluett átlag*(1-gradiens), harmadik sor: Hurst-exponens) erre a két csoportosításra volt a legnagyobb értékű.

20. táblázat Klaszterezési módszerek összehasonlító számadatai

	Módszer 1	**Módszer 2**	**Módszer 3**
sziluett átlag	0,59	0,72	0,72
sziluett átlag*(1-gradiens)	0,48	0,62	0,62
Hurst-exponens	0,52	1	1

A szintetikus adatrendszeren történő vizsgálatot követően a k-means, MFV és WMFV klaszterezést a nílusi adatrendszeren is futtattam. A klaszterezések bemenő változói az évenkénti minimum, maximum, illetve maximum-minimum különbségi vízmélység adatok voltak. A csoportosítási módszerek a 20. ábrán látható eredményt adták.



20. ábra. Klaszterezési eljárások eredménye a Nílus-vízszintmérések adatrendszerén

Az eredményekből látszik, hogy a korai évek csoportjainál egy természetes ingadozás figyelhető meg, az 1600-as években megjelenik egy új klaszter, amely egybeesik az adatsor manuális értelmezései során tapasztaltakéval [37], [34].

A csoportosítási eljárások összehasonlításának mérőszámait a 21. táblázat tartalmazza. Ezek a szintetikus adatrendszerben alkalmazottal megegyezőek voltak: első sor: sziluett átlag, második sor: sziluett átlag*(1-gradiens), harmadik sor: Hurst-exponens.

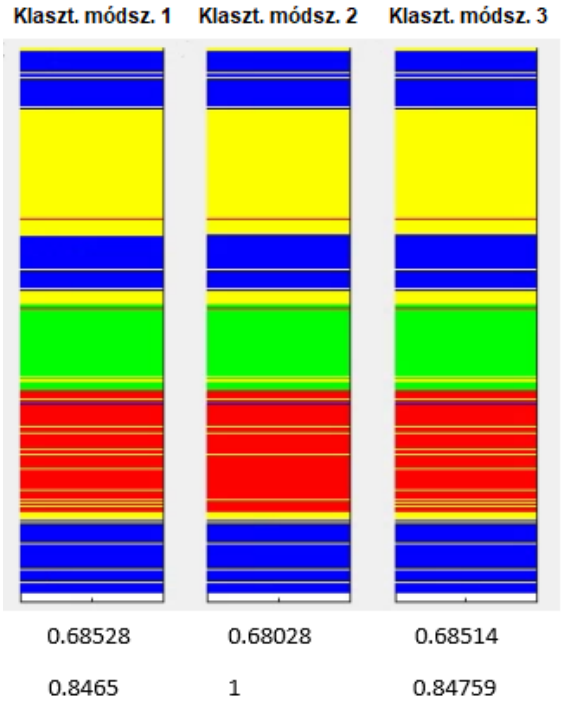
21. Táblázat Klaszterelemzés eredményeinek számszerű összehasonlítása

	k-means	**MFV**	**WMFV**
sziluett átlag	0,5082	0,4109	0,51358
sziluett átlag*(1-gradiens)	0,32662	0,25649	0,33008
Hurst-exponens	0,87791	0,7792	0,88664

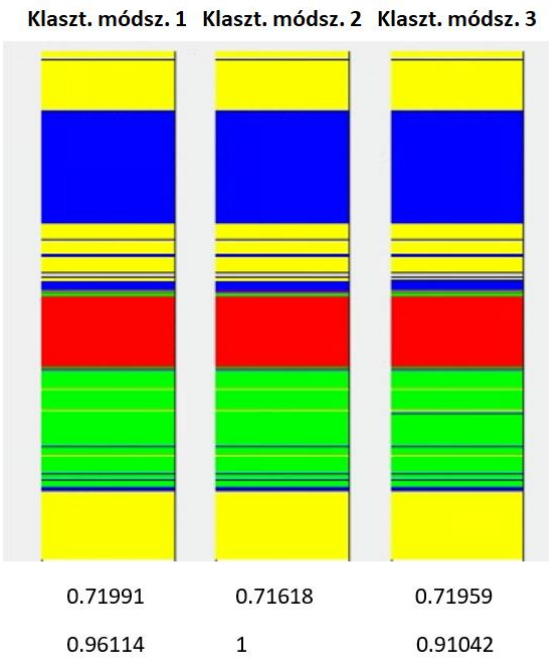
A 21. táblázatból kiolvasható, hogy a WMFV klaszterezési eljárást mutatták a mérőszámok a legjobbnak (mindhárom mérőszám esetében a harmadik oszlop értéke volt a legnagyobb).

3.4.8 A Hurst exponensen alapuló mérőszám további összehasonlító vizsgálata

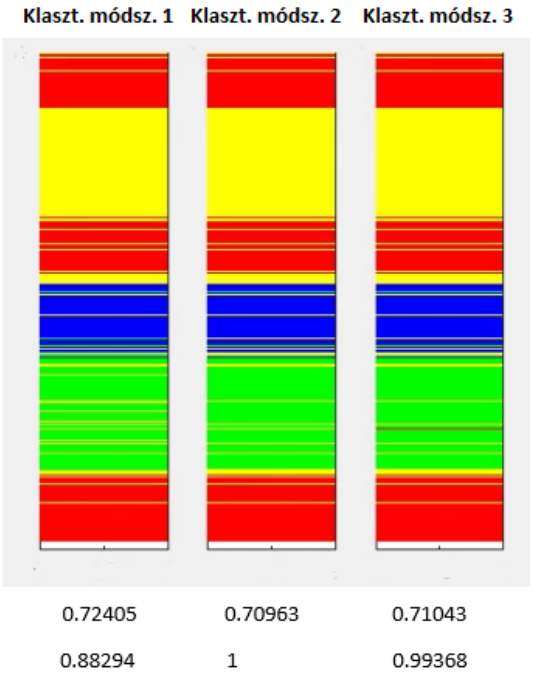
A fenti eseten túl további futtatások esetében is vizsgáltam a mérőszámokat egyazon (kiinduló) szintetikus adatrendszeren, az előbbihez hasonlóan véletlenszerű zajjal, ezen futtatások példa eredményeit a 21-23. ábrák tartalmazzák.



21. ábra. Klaszterelemzést jellemző mérőszám vizsgálata (első példa)



22. ábra. Klaszterelemzést jellemző mérőszám vizsgálata (második példa)



23. ábra. Klaszterelemzést jellemző mérőszám vizsgálata (harmadik példa)

A 21., 22., 23. ábrák mindegyike esetében az adott klaszterezési eredmény vizuális reprezentációja alatt fentebb a sziluett érték (sziluett átlag), lentebb a Hurst-exponens adta mérőszám értéke látható. A példák mindegyike esetében az olvasható le az eredményekből,

hogy ezekben az esetekben az utóbbi mutatta meg helyesen, hogy melyik volt a legjobb klaszterezési kimenet. Megjegyzendő tulajdonság itt az, hogy a Hurst-exponens értékét minden esetben a legnagyobb értékhez skáláztam.

3.4.9 Tézis

5. Tézis: A Hurst-exponenst arányszámként alkalmaztam a klaszterelemzés jóságának jellemzésére. Az exponens értékét mind a klaszterindexek vektorára, mind pedig az eredeti adatsorra kiszámítva, az így kapott arány – az átlagos sziluett érték és más jósági jellemzőkhöz hasonlóan – felhasználható mérőszámként az eredmények minőségének jellemzésére. A validáláshoz használt adatrendszerekre a Hurst-exponensen alapuló mérőszám összehasonlítható eredményt produkált és nem maradt el a sziluett átlagtól, mint jósági jellemzőtől. A fenti megállapításokat in situ hidrológiai adatsoron és fúrólyuk-geofizikai szelvények felhasználásával is alátámasztottam.

Kapcsolódó saját publikáció: [S8]

3.5 A WMFV klaszterelemzési eljárás súlyparaméterének adaptívvá tétele

A nílusi adatrendszeren végzett vizsgálatok során megalkottam a WMFV klaszterezési eljárás az idáig konstans $\mathbf{w}=[w_1 \;\; w_2]^T$ súlyainak dinamikus változatát az alábbi formában. Az eddigi két komponens (m_{e1} *és* m_{e2}) helyett itt négyet alkalmaztam: m_{e1}, m_{e2}, MFV, ill. a k-means eljárás adta klaszterközéppont. A négy változó közül mindet rendre az adatablak mátrixába téve, a legkisebb és második legkisebb Mahalanobis-távolságú változó értékeként meghatároztam $P1_{v1}$*és* $P1_{v2}$ értékeket. Ezekből P1 értékét súlyozott átlaggal képeztem, ahol az első ($P1_{v1}$) komponens w_{p1v1} súlyát a komponenst az ablak adatmátrixába helyezve a helyettesítés előtti adatablakkal való

$$\text{átlag}(1/\text{Mahalanobis-távolság}) * \text{robusztus\_kovariancia} \qquad (32)$$

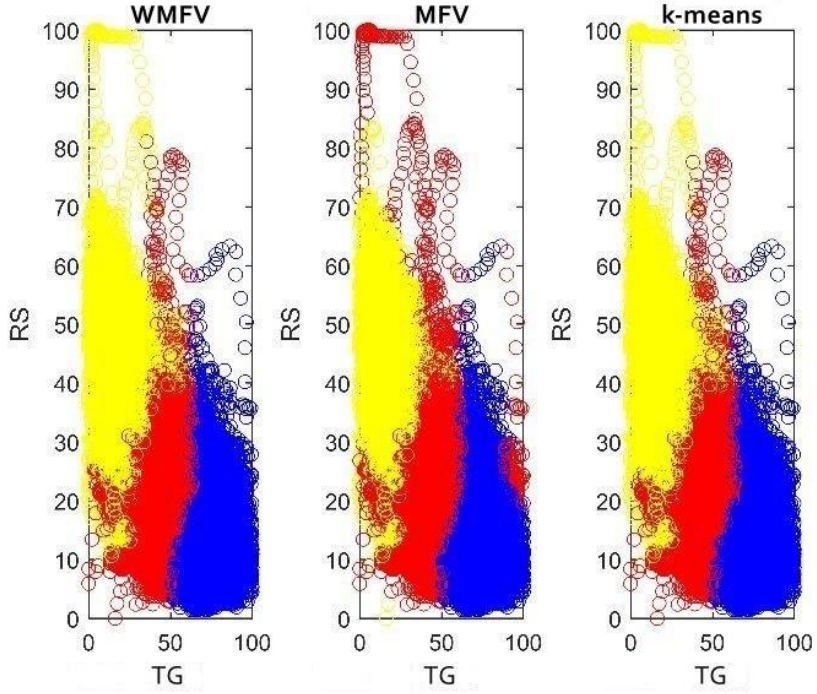
érték adta. $P1_{v2}$ súlya esetében hasonlóképp jártam el. (A robusztus kovarianciát a [38] publikáció szerint használtam.)

Ezt követően P2 értékét határoztam meg, itt mind a négy kiinduló változót az adatablak közepére téve, a maximális robusztus kovariancia értékű változó értékét véve.

Végül az adatablak középpontjának végleges értékeként P1 és P2 súlyozott átlagát számítottam ki $\mathbf{w}=[w_1 \;\; w_2]^T$ súlyaikkal, ahol $w_1$értéke: 1/(P1 Mahalanobis-távolsága $[P1_{v1} P1_{v2}]$ vektortól), w_2 értéke hasonlóképp, 1/(P2 Mahalanobis-távolsága $[P1_{v1} P1_{v2}]$ vektortól). A fent ismertetett,

az adatsoron végzett vizsgálatok összehasonlító eredményei már a WMFV módszer ezen változatával számítottak.

Az előbbieken túl egy további adatrendszeren, Debrecen környéki, nagy földtani és hidrogeológiai változékonyságot mutató mélyfúrási geofizikai adatokon is vizsgáltam módszerem [S9].



24. ábra TG-RS klaszterelemzési eredmények megjelenítése egy Debrecen környéki fúrás adatrendszere esetén

A 24. ábrán látható vizuális reprezentációban megfigyelhető a példaként bemutatott természetes gamma-sugárzási adatok (TG) ill. sekély-fajlagos ellenállás (RS) esetében, hogy a WMFV módszer sokkal jobban szeparálta a 3 klaszter adatait, mint a hagyományos MFV eljárás. A k-means klaszterezés ez esetben vizuálisan hasonló eredményt produkált. Mindez mind 3, mind 4-es klaszterszám esetén fennállt, azonban mivel 4 klaszter esetében az átlagos sziluett érték mindhárom módszer esetében ~0,29-nek adódott, míg ~0,42 volt 3 klaszter esetében, az utóbbi tekinthető az ideális klaszterszámnak az adatrendszeren, így ezt vettem alapul.

Megjegyzendő a fentieken túl, hogy ezen az adatrendszeren módszerem, a k-means, és a tiszta MFV klaszterezés mellett a SOM (Self Organizing Map) eljárás [31] is futtatásra került. Utóbbi tanítása elsőként 900 neuronnal, 30x30-as térképmérettel, majd 4 neuronnal, 2x2-es

térképmérettel történt meg, a módszerrel és a bemenetként használt adatok területével foglalkozó földtudományi szakértők által [S9].

A módszerek által kapott eredményekben mind a k-means, mind a SOM eredmények alulreprezentálták a homok arányát a szakértői jellemzések szerint, ellentétben az MFV, és WMFV eljárásokkal.

3.6 Az eredmények összegzése

A fentiek összegzéseként kijelenthető, hogy a hisztogramalapú súlyozott medián eljárás klaszterezésre megalkotott, ismertetett változata a vizsgált egyéb módszerekkel összehasonlítva is jó eredményt produkált heterogén föltudományi adatokon. Ezt nem csupán bevett metrikákkal igazoltam, hanem egy általam megalkotott mérőszám segítségével is, amelyet külön is összehasonlító validációnak vetettem alá, azt igazolandó, hogy az a vizsgált metrikákkal hasonlóan eredményesen alkalmazható. Az előbbieken túl a klaszterezési eljárás egy továbbfejlesztett, adaptív súlyozásos változatát is bemutattam.

3.6.1 Tézis

6. Tézis: A robusztus klaszterelemzési eljárás (WMFV) továbbfejlesztése során az alábbi eredményeket értem el:

Megállapítottam, hogy az eljárásban szereplő fő súlyok adaptív módon meghatározhatók. Megállapítottam, hogy a súlyozott medián eljárás és az MFV módszer kombinációján alapuló eljárás mind adaptív súlyos változatában mind eredeti formájában a hagyományos MFV, illetve k-means eljárásokkal összemérhető eredményt nyújt. Ezt igazoltam heterogén földtudományi adatrendszereken többféle metrika (sziluett módszer, gradiens, korreláció), ill. fúrólyuk-geofizikai szelvények független értelmezési eredményeinek felhasználásával. A módszert kiterjesztettem fúrólyukgeofizikai adatok együttes (több fúrásra kiterjedő) értékelésére.

Kapcsolódó saját publikációk: [S7, S8, S9]

Új tudományos eredmények

1. Tézis:

Kidolgoztam egy kizárólag hanginformációkat felhasználó, szabályalapú osztályozási módszert, amely képes a fejmozgás mértékét kellő pontossággal becsülni. A módszer a betanításban nem szereplő alanyok mozgásának tartományát is képes becsülni. Bebizonyítottam, hogy képezhetők olyan hanginformációs paraméterek, amelyek mind a mozgásmérték osztályozását, mind pedig a hang és a fejmozgás közötti korrelációt tekintve jobb eredményt adnak, mint a hagyományosan alkalmazott hanginformációs jellemzők. Igazoltam, hogy az általam megalkotott osztályozási eljárás jelen struktúrájában megfelelő megoldást ad. A módszer az általam vizsgált esetekben jobb megoldást ad, mint a dimenziócsökkentéssel támogatott változat, illetve a gyakorlatban hagyományosan alkalmazott eljárások.

Kapcsolódó saját publikációk: [S1, S2, S3]

2. Tézis:

Megalkottam a mozgásmérték osztályozási eljárás hatékonyságának növelése céljából egy a kieső értékeket hatékonyan kiszűrni képes, Mahalanobis és euklideszi távolság együttes használatával operáló outlier-detektálási eljárást.

Kapcsolódó saját publikáció: [S3]

3. Tézis:

Kifejlesztettem egy új, hisztogramalapú zajszűrési eljárást, amely a leggyakoribb érték módszerét (Most Frequent Value-t, röviden MFV-t) felhasználja topográfiai adatok szűrésére. A fenti módszer részben az adatok dinamikus mértékű, hisztogramalapú előszűrésével, MFV-t, illetve további súlyozásokat felhasználva a natív MFV-vel és a hagyományos mediánszűréssel összemérhető eredményt produkál. Az eljárás heterogén adatrendszereken szignifikánsan jobb zajcsökkenést produkál a hagyományos mediánszűréshez képest. Megállapítottam, hogy a hisztogramalapú szűrési eljárásban kiszámított részeredmények abszolút különbsége nagy megbízhatósággal képes előre jelezni, hogy a hisztogramalapú előszűrés jó eredménnyel alkalmazható-e az adott adatablakban vagy ellenkezőleg. Ezen

megállapítást az L_1-norma alapján számított hiba és a részeredmények különbségének számértéke közötti 0,99 értékű korrelációs együttható támasztja alá. A megfigyelést a szűrési eljárás eredményének súlyozása során felhasználtam módszerem hatékonyságának és robusztusságának növelésére, az adott ablakbeli MFV és hisztogramalapú szűrési eredmény értéke közötti súlyozás formájában.

Kapcsolódó saját publikáció: [S4]

4. Tézis:

Kifejlesztettem egy leggyakoribb érték (MFV) - alapú klaszterelemzési módszert, ahol a kezdeti centroidokat hisztogram-alapú szűréssel határoztam meg. A módszert mélyfúrási geofizikai adatokra alkalmaztam és meghatároztam a klaszterek mélységi változását. A kidolgozott súlyozási eljárással hatékonyan megtisztítottam az eredményt a kieső értékektől. Ezen felül bemutattam, hogy a módszer eltérő zajszennyezettség és adathiány mértékek esetében is robusztus, megfelelő minőségi jellemzőkkel bíró csoportosítási alternatívát jelent. Az eljárást több fúrás adatrendszereinek együttes feldolgozására is alkalmassá tettem. A robusztus eljárás javítja a klaszterelemzés zajérzékenységét, biztosítja a kőzetrétegek megfelelő felbontását, segíti a nyersanyagok azonosításának hatékonyságát és a mérési adatok földtani-geofizikai értelmezését.

Kapcsolódó saját publikáció: [S7]

5. Tézis:

A Hurst-exponenst arányszámként alkalmaztam a klaszterelemzés jóságának jellemzésére. Az exponens értékét mind a klaszterindexek vektorára, mind pedig az eredeti adatsorra kiszámítva, az így kapott arány – az átlagos sziluett érték és más jósági jellemzőkhöz hasonlóan – felhasználható mérőszámként az eredmények minőségének jellemzésére. A validáláshoz használt adatrendszerekre a Hurst-exponensen alapuló mérőszám összehasonlítható eredményt produkált és nem maradt el a sziluett átlagtól, mint jósági jellemzőtől. A fenti megállapításokat in situ hidrológiai adatsoron és fúrólyuk-geofizikai szelvények felhasználásával is alátámasztottam.

Kapcsolódó saját publikáció: [S8]

6. Tézis:

A robusztus klaszterelemzési eljárás (WMFV) továbbfejlesztése során az alábbi eredményeket értem el:

Megállapítottam, hogy az eljárásban szereplő fő súlyok adaptív módon meghatározhatók. Megállapítottam, hogy a súlyozott medián eljárás és az MFV módszer kombinációján alapuló eljárás mind adaptív súlyos változatában mind eredeti formájában a hagyományos MFV, illetve k-means eljárásokkal összemérhető eredményt nyújt. Ezt igazoltam heterogén földtudományi adatrendszereken többféle metrika (sziluett módszer, gradiens, korreláció), ill. fúrólyuk-geofizikai szelvények független értelmezési eredményeinek felhasználásával. A módszert kiterjesztettem fúrólyukgeofizikai adatok együttes (több fúrásra kiterjedő) értékelésére.

Kapcsolódó saját publikációk: [S7, S8, S9]

New scientific results

Thesis 1: I have developed a rule-based classification method utilizing exclusively acoustic information, capable of estimating head movement metrics with sufficient accuracy. The method is also capable of estimating the movement range of subjects not included in the training set. I have demonstrated that acoustic parameters can be derived which yield better results—both in terms of motion scale classification and the correlation between audio and head movement—than traditionally applied acoustic features. I have verified that the proposed classification procedure provides an appropriate solution in its current structure. In the cases examined, the method outperforms both the version supported by dimensionality reduction and the procedures conventionally used in practice.

Related publications: [S1, S2, S3]

Thesis 2: I have created an outlier detection procedure operating with the combined use of Mahalanobis and Euclidean distances, specifically designed to increase the efficiency of the motion scale classification process.

Related publication: [S3]

Thesis 3:I have developed a new histogram-based noise filtering procedure that utilizes the Most Frequent Value (MFV) method for filtering topographical data. By employing dynamic histogram-based pre-filtering, MFV, and additional weighting, the method produces results comparable to native MFV and traditional median filtering. On heterogeneous datasets, the procedure yields significantly better noise reduction compared to traditional median filtering. I have established that the absolute difference between sub-results calculated in the histogram-based filtering can predict with high reliability whether the pre-filtering will be effective in a given data window. This finding is supported by a correlation coefficient of 0.99 between the numerical value of the difference in sub-results and the error calculated based on the L_1-norm. I utilized this observation to increase the efficiency and robustness of the method by weighting the filtering result between the MFV of the given window and the histogram-based filtering result.

Related publication: [S4]

Thesis 4: I have developed an MFV-based cluster analysis method, where initial centroids are determined via histogram-based filtering. I applied the method to borehole geophysical data and determined the depth-wise variation of clusters. Using the developed weighting procedure, I effectively cleared the results of outliers. Furthermore, I have demonstrated that the method is robust even under varying levels of noise and data gaps, providing a clustering alternative with appropriate quality characteristics. I have extended the procedure to enable the joint processing of datasets from multiple boreholes. This robust procedure improves the noise sensitivity of cluster analysis, ensures proper resolution of rock layers, and enhances the efficiency of raw material identification and the geological-geophysical interpretation of measurement data.

Related publication: [S7]

Thesis 5: I have applied the Hurst exponent as a ratio to characterize the quality of cluster analysis. By calculating the exponent for both the vector of cluster indices and the original dataset, the resulting ratio—similarly to the average silhouette value and other goodness-of-fit metrics—can be used as a measure to characterize the quality of the results. For the datasets used for validation, the Hurst exponent-based metric produced comparable results and did not underperform relative to the average silhouette value. I have supported these findings using both in situ hydrological datasets and well logs.

Related publication: [S8]

Thesis 6: During the further development of the robust cluster analysis procedure (WMFV), I achieved the following results:

I established that the primary weights in the procedure can be determined in an adaptive manner. I found that the method based on the combination of the weighted median procedure and the MFV method—both in its adaptive-weight version and its original form—provides results comparable to traditional MFV and k-means procedures. I have verified this on heterogeneous geoscientific datasets using various metrics (silhouette method, gradient, correlation) and independent interpretation results of well logs. I extended the method to the joint evaluation of multi-borehole geophysical data.

Related publications: [S7,S8,S9]

Összefoglalás

A disszertációban az adat- és jelfeldolgozás témakörében végzett módszerfejlesztési kutatásaim során elért eredményeket mutattam be. Elsőként egy virtuális ügynökök mozgásának természetességének növelésére hivatott, fejmozgás mértékek pusztán hanginformációk alapján történő becslésén alapuló osztályozási és mozgásmérték skálázási módszert ismertettem, amely akár annak megállapításában is segítség lehet, hogy egy adott videó valós, vagy generált beszélőt tartalmaz. Az osztályozási módszer megalkotása során többek között megállapításokat tettem derivált hangjellemzők és mozgásmérték közötti erős korrelációs kapcsolatokra is, amely az eljárás egyik továbbfejlesztési iránya lehet az egyéb ismertetett lehetőségeken túl. Egy az osztályozási eljárás eredményeinek javítása céljával kifejlesztett egyedi outlier-detektálási eljárást is ismertettem a disszertáció ezen fejezetében.

Az értekezés második pontjában az adatok súlyozott hisztogramszűréssel való előfeldolgozásán és további súlyozásokon alapuló – főként digitális domborzati adatok kezelésére irányuló – zajszűrési eljárást mutattam be, amely az eredmények alapján jobb eredményt produkál a vizsgált esetekben a hagyományos mediánszűrésnél, de felveszi a versenyt a számos tudományterületen sikeresen alkalmazott leggyakoribb érték (MFV) robosztus módszerével is.

A disszertáció harmadik pontjában az ismertetett zajszűrési eljárás alapötletét és tapasztalatait felhasználva megalkotott klaszterezési módszertant mutattam be, amely heterogén földtudományi adatokon (például terepi egy- ill. kétdimenziós mélyfúrási adatok, vízmélység adatok) vizsgálva is bizonyította hatékonyságát. Ezen felül a klaszterezési eljárások kvantitatív minősítésére vizsgálataim szerint sikeresen alkalmazható általános mérőszámot is ismertettem munkámban.

Summary

In my dissertation, I presented the results of my methodological development research in the field of data and signal processing.

First, a classification and motion-scaling framework is introduced to enhance the perceived naturalness of virtual agents, based on the estimation of head kinematic parameters solely from acoustic input. Beyond animation, this procedure can also assist in determining whether a specific video features a real or a generated speaker.

My analysis established strong correlations between derivative acoustic features and kinematic parameters, offering a promising trajectory for future research. Furthermore, I developed a specialized outlier detection procedure to optimize classification accuracy.

The second part of the thesis introduced a noise filtering procedure—primarily aimed at handling digital terrain data—based on pre-processing data via weighted histogram filtering and additional weighting techniques. Based on the results, this method outperforms traditional median filtering in the cases examined and comparable to the Most Frequent Value (MFV) robust method, which is successfully applied across numerous scientific disciplines.

In the third section, utilizing the core concepts and experiences of the aforementioned noise filtering procedure, I presented a clustering methodology. This approach proved its efficiency when tested on heterogeneous geoscientific data, such as in situ 1D and 2D well logging datasets and long term water level measurements. Furthermore, I introduced a general metric that, according to my research, can be successfully applied to the quantitative qualification of clustering procedures.

Hivatkozások

[1] Abdalla, H. B. (2022). A brief survey on big data: technologies, terminologies and data-intensive applications. *Journal of Big Data, 9(1)*, 107.

[2] Al-Mudhafar, W. J., & Bondarenko, M. A. (2015). Integrating K-means clustering analysis and Generalized Additive Model for efficient reservoir characterization. *77th EAGE Conference and Exhibition 2015*, (old.: 1-6.). Forrás: https://doi.org/10.3997/2214-4609.201413024

[3] Al-Mudhafar, W. J., Al Lawe, E. M., & Noshi, C. I. (2019). Clustering analysis for improved characterization of carbonate reservoirs in a Southern Iraqi oil field. *Offshore Technology Conference.* Houston, USA. Forrás: https://doi.org/10.4043/29269-MS

[4] Baudat, G., & Anouar, F. (2000). Generalized discriminant analysis using a kernel approach. *Neural computation, 12(10)*, 2385-2404.

[5] Ben, Y., A., S. H., & Braude, D. A. (2013). Articulatory features for speech-driven head motion synthesis. *Proceedings of Interspeech.* Lyon, Franciaország.

[6] Calinszki, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods, 3(1)*, 1-27.

[7] Chen, F. F., Yang, Y. S., Pervukhina, M., Clennell, B. M., & Taylor, J. A. (2016). Clustering analysis for porous media: An application to a dolomitic limestone. *J. Pet. Sci. Eng., 146.*, 770-776. Forrás: https://doi.org/10.1016/j.petrol.2016.07.031

[8] Cheng, Y., & Church, G. M. (2000).). Biclustering of expression data. *ISMB, 8*, 93-103.

[9] Cowgill, M. C., Harvey, R. J., & Watson, L. T. (1999). A genetic algorithm approach to cluster analysis. *Computers & Mathematics with Applications, 37.*, 99-108. Forrás: https://doi.org/10.1016/S0898-1221(99)00090-5

[10] Davies, L. D., & Bouldin, W. D. (1979). A cluster separation measure. *IEEE transactions on pattern analysis and machine intelligence, 2*, 224- 227.

[11] Deng, Z., Narayanan, S., Busso, C., & U., N. (2004). Audio-based head motion synthesis for avatar-based telepresence systems. *Proceedings of the 2004 ACM SIGMM workshop on Effective telepresence*, (old.: 24-30.).

[12] Dobróka, M., Gyulai, Á., Ormos, T., Csókás, J., & Dresen, L. (1991). Joint inversion of seismic and geoelectric data recorded in an underground coal mine. *Geophysical Prospecting, 39 (5)*, 643-665. Forrás: https://doi.org/10.1111/j.1365-2478.1991.tb00334.x

[13] Dobróka, T. (2021). An MFV-based image processing filter and its use in seismic tomograph. *Acta Geodaetica et Geophysica, 56(3).*

[14] Dong, S., Wang, P., & Abbas, K. (2021). A survey on deep learning and its applications. *Computer Science Review, 40*, 100379.

[15] Filzmoser, P., Garrett, R. G., & Reimann, C. (2005). Multivariate outlier detection in exploration geochemistry. *Computers & Geosciences, 31.*, 579-587. Forrás: https://doi.org/10.1016/j.cageo.2004.11.013

[16] Fujita, A., Takahashi, Y. D., & Patriota, G. A. (2014). A nonparametric method to estimate the number of clusters. *Computational Statistics & Data Analysis, 73*, 27-39.

[17] Gallegos, M. T., & Ritter, G. (2005). A robust method for cluster analysis. *The Annals of Statistics, 33.*, 347-380. Forrás: https://doi.org/10.1214/009053604000000940

[18] Getz, G., E., L., & Domany, E. (2000). Coupled two-way clustering analysis of gene microarray data. *Proceedings of the National Academy of Sciences, 97(22)*, 12079-12084.

[19] Greenwood, D., Laycock, S., & Matthews, I. (2017). Predicting head pose from speech with a conditional variational autoencoder. *Interspeech 2017*, (old.: 3991-3995.).

[20] Grimm, M., Neumann, U., Busso, C., Deng, Z., & S., N. (2007). Rigid head motion in expressive speech animation: Analysis and synthesis. *IEEE Transactions on Audio, Speech, and Language Processing, 3*, 1075-1086.

[21] Grimm, M., Neumann, U., Busso, C., Z., D., & S., N. (2005). Natural head motion synthesis driven by acoustic prosodic features. *Journal of Visualization and Computer Animation*, (3-5) 283-290.

[22] Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: k-means clustering algorithm. *J. R. Stat. Soc. Ser. C Appl. Stat., 28.*, 100-108. Forrás: https://doi.org/10.2307/2346830

[23] Hempkins, W. B. (1978). Multivariate statistical analysis in formation evaluation. *SPE California Regional Meeting.* San Francisco, USA. doi:7144-MS

[24] Hofer, G., & Shimodaira, H. (2007).Automatic head motion prediction from speech data. *Interspeech 2007*, (old.: 722-725.).

[25] Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology, 24(6),* 417-441. doi:10.1037/h0071325

[26] Huang, T. S., Yang, G. J., & Tang, G. Y. (1979). A fast two-dimensional median filtering algorithm. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 27 (1)*, 13–18. Forrás: https://doi.org/10.1109/TASSP.1979.1163188

[27] Hurst, H. (1951). Long term storage capacity of reservoirs. *Trans Am Soc Eng, 116*, 770–799.

[28] Jarzyna, J. e. (2017). *Shale gas in Poland: Advances in natural gas emerging technologies*. (H. Al-Megren, & R. Altamimi) London: IntechOpen. Forrás: https://doi.org/10.5772/67301

[29] Ji, X. e. (2022). Eamm: One-shot emotional talking face via audio-based emotion-aware motion model. *ACM SIGGRAPH 2022 Conference Proceedings.*

[30] Kim, H., Garrido, P., Tewari, A., Xu, W., Thies, J., Nießner, M., Theobalt, C. (2018). Deep video portraits. *ACM Transactions on Graphics (TOG), 37(4)*, 1-14.

[31] Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological cybernetics, 43(1)*, 59-69.

[32] Liu, X., Yin, J., Feng, Z., Dong, J., & Lu., W. (2007). Orthogonal neighborhood preserving embedding for face recognition. *Image Processing, 2007. ICIP 2007. IEEE nemzetközi konferencia, 1*, old.: 133-136.

[33] Lu, Y., Chai, J., & Cao, X. (2021). Live speech portraits: real-time photorealistic talking-head animation. *ACM Transactions on Graphics (TOG), 40(6)*, 1-17.

[34] Lüdecke, H. J., Müller-Plath, G., Wallace, M. G., & Lüning, S. (2021). Decadal and multidecadal natural variability of African rainfall. *Journal of Hydrology: Regional Studies*, 34.

[35] Matthews, I., S., L., & Greenwood, D. (2018). Joint learning of facial expression and head pose from speech. *15*, 2484-2488.

[36] Nasraoui, O., & N'Cir, C.-E. B. (2019). *Clustering methods for big data analytics: Techniques, Toolboxes and Applications.* Springer Cham. Forrás: https://doi.org/10.1007/978-3-319-97864-2

[37] O'Sullivan, P. E., Moyeed, R., Cooper, M. C., & Nicholson, M. J. (2002). Comparison between instrumental, observational and high resolution proxy sedimentary records of Late Holocene climatic change—a discussion of possibilities. *Quaternary International, 88(1),* 27–44.

[38] Olive, D. J., & Hawkins, D. M. (2010). Robust multivariate location and dispersion. *Preprint*.

[39] Rousseeuw P., J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics, 20,* 53-65.

[40] Roweis, S. T. (2002). Automatic alignment of hidden representations. *Sixteenth Annual Conference on Neural Information Processing Systems, 15,* old.: 841-848. Vancouver, Kanada.

[41] Scott, D. W. (1979). On optimal and data-based histograms. *Biometrika, 66(3),* 605–610.

[42] Scott, D. W. (1992). *Multivariate density estimation: theory, practice, and visualization.* New York: John Wiley & Sons.

[43] Sfidari, E., Kadkhodaie-Ilkhchi, A., & Najjari, S. (2012). Comparison of intelligent and statistical clustering approaches to predicting total organic carbon using intelligent systems. *J. Pet. Sci. Eng., 86-87.*, 190-205. Forrás: https://doi.org/10.1016/j.petrol.2012

[44] Silva, D. V., & Tenenbaum, B. J. (2003). Global versus local methods in nonlinear dimensionality reduction. *Advances in neural information processing systems, 120,* 721-728.

[45] Spearman, C. (1904). 'General intelligence', objectively determined and measured. *The American Journal of Psychology, 15(2),* 201–293. doi:10.2307/1412107

[46] Steiner, F. (1991). *The Most Frequent Value. Introduction to Modern Conception Statistics.* Budapest: Akadémiai Kiadó.

[47] Stone, D. C. (1995). Application of median filtering to noisy data. *Canadian Journal of Chemistry, 73(10),* 1573-1581. Forrás: https://doi.org/10.1139/v95-195

[48] Szabó, N. P., Balogh, G. P., & Stickel, J. (2018). Most frequent value-based factor analysis of direct-push logging data. *Geophysical Prospecting, 66.* Forrás: https://doi.org/10.1111/1365-2478.12573

[49] Szabó, N. P., Braun, B. A., Abdelrahman, M. M., & Dobróka, M. (2021). Improved well logs clustering algorithm for shale gas identification and formation evaluation. *Acta Geod Geophys, 56.*, 711-729. Forrás: https://doi.org/10.1007/s40328-021-00358-0

[50] Szabó, N. P., Dobróka, M., & Kavanda, R. (2013). Cluster analysis assisted float-encoded genetic algorithm for a more automated characterization of hydrocarbon reservoirs. *Intelligent Control and Automation, 4.*, 362-370. Forrás: https://doi.org/10.4236/ica.2013.44043

[51] Szabó, N. P., Nehéz, K., Hornyák, O., Piller, I., Deák, C., Hanzelik, P. P., . . . Ott, K. (2019). Cluster analysis of core measurements using heterogeneous data sources: An application to complex Miocene reservoirs. *J. Pet. Sci. Eng., 178.*, 575-585. Forrás: https://doi.org/10.1016/j.petrol.2019.03.067

[52] Szabó, N. P., Valadez-Vergara, R., Tapdigli, S., Ugochukwu, A., Szabó, I., & Dobróka, M. (2021). Factor analysis of well logs for total organic carbon estimation in unconventional reservoirs. *Energies, 14.*, 5978. Forrás: https://doi.org/10.3390/en14185978

[53] Szabó, N., & Balogh, G. P. (2016). Most frequent value based factor analysis of engineering geophysical sounding logs. *78th EAGE Conference and Exhibition 2016*. Houten, Holland: European Association of Geoscientists and Engineers (EAGE). Forrás: https://doi.org/10.3997/2214-4609.201600796

[54] Szűcs, P., & Zákányi, B. (2007). Applying most frequent value (MFV) in hydrogeological modelling. *Mérnökgeológia-Kőzetmechanika 2007*, (old.: 161-174.).

[55] Szűcs, P., Dobróka, M., Turai, E., Szarka, L., Ilyés, C., Eid, M. H., & Szabó, N. P. (2024). Combined inversion and statistical workflow for advanced temporal analysis of the Nile River's long term water level records. *Journal of Hydrology, 630*, 130693.

[56] Tavakoli, V. (2018). *Geological Core Analysis: Application to Reservoir Characterization.* Springer Cham. Forrás: https://doi.org/10.1007/978-3-319-78027-6

[57] Teh, W., Willhite, G. P., & Doveton, J. H. (2012). Improved reservoir characterization using petrophysical classifiers within electrofacies. *SPE Improved Oil Recovery Symposium*. Tulsa, USA. Forrás: https://doi.org/10.2118/154341-MS

[58] Tibshirani, R., Walther, G., & T., H. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(2)*, 411-423.

[59] Toussoun, O. (1925). *Mémoire sur l'Histoire du Nil. Mémoire Institut d'Egypte 8e.* Cairo: Le Caire Imprimerie de L'Institut Francais.

[60] Völgyesi, L., & Tóth, G. (2021). Improvement of QDaedalus measurements with continuous detection of environmental parameters. *Acta Geod Geophys, 56*, 607-622. doi:10.1007/s40328-021-00359-z

[61] Wang, W., Wang, Z., Leung, J. Y., Kong, C., & Jiang, Q. (2022). Petrophysical rock typing based on deep learning network and hierarchical clustering for volcanic reservoirs. *J. Pet. Sci. Eng., 210*, 110017. Forrás: https://doi.org/10.1016/j.petrol.2021.110017

[62] Yang, H., Pan, H., Ma, H., Konaté, A. A., Y. J., & Guo, B. (2016). Performance of the synergetic wavelet transform and modified K-means clustering in lithology classification using nuclear log. *J. Pet. Sci. Eng., 144.*, 1-9. Forrás: https://doi.org/10.1016/j.petrol.2

[63] Yehia, H. C., Kuratate, T., & E., V.-B. (2002). Linking facial animation, head motion and speech acoustics. *Journal of phonetics, 30(3)*, 555-568.

[64] Zhang, J. (2017). Most frequent value statistics and distribution of 7Li abundance observations. *Monthly Notices of the Royal Astronomical Society, 468(4)*, 5014–5019. Forrás: https://doi.org/10.1093/mnras/stx627

[65] Zhang, K., Lin, N., Fu, C., Zhang, D., Jin, X., & Zhang, C. (2019). Reservoir characterisation method with multi-component seismic data by unsupervised learning and colour feature blending. *Exploration Geophysics, 50.*, 269-280. Forrás: https://doi.org/10.1080/0812398

[66] Zhou, Y. H. (2020). MakeltTalk: speaker-aware talking-head animation. *ACM Transactions on Graphics (TOG), 39(6)*, 1-15.

Saját publikációk az értekezés témakörében

[S1] Czap, L., & Kilik, R. (2015). Automatic gesture generation. *Production Systems and Information Engineering, 7*, 5–14.

[S2] Kilik, R. (2023). Estimation of the Extent of Vertical Head Movements From Sound Alone. *Acta Marisiensis. Seria Technologica, 20(1)*, 1-9.

[S3] Roland, K. (2016.). Fejmozgás mértékek hanganyagából történő megállapításának kezdeti eredményei. *MultiScience - XXX. microCAD International Multidisciplinary Scientific Conference University of Miskolc*. Magyarország. doi:10.26649/musci.2016.064

[S4] (Q3 kvartilis) Kilik, R. (2021). Histogram-based weighted median filtering used for noise reduction of digital elevation model data. *Acta Geodaetica et Geophysica, 56(4)*, 743-764.

[S5] Kilik, R. (2020). Applying iterative median-based filtering method for enhancing catchment area models. *Műszaki Földtudományi Közlemények, 89(1)*, 357-363.

[S6] Pecsmány, P., & Kilik, R. (2019). A leggyakoribb érték módszerének alkalmazási lehetősége a digitális domborzatmodellek zajszűrésére. *Miskolci Egyetem.*

[S7] (Q1 kvartilis) Szabó, N. P., Kilik, R., & Dobróka, M. (2023). Robust reservoir identification by multi-well cluster analysis of wireline logging data. *Heliyon, 9(5)*.

[S8] Kilik, R., Ilyés, Cs. (2023) Klaszter és leíró statiszika elemzés közel 1500 éves vízszint adatsoron. *Új eredmények a műszaki föld-és környezetttudományban*, 318-328.

[S9] (Q1 kvartilis) Mohammed, M. A., Szabó, N. P., Kilik, R., & Szűcs, P. (2024). Examining innovative unsupervised learning techniques for automated characterization of complex groundwater systems. *Results in Engineering, 23*, 102594.

Függelék

1. táblázat Kinyert hanginformációs jellemzők

P_1		$\overline{autokorreláció_{pitch}}$
P_2		$var\big(FFT(intenzitás)\big)$
P_3		$\frac{\sum_{i=2}^{N}\lvert ampl_i - ampl_{i-1}\rvert}{Képkockákszáma}$ ahol: $ampl_k = \frac{\sum_{j=k-r}^{k+r}\lvert hang_j\rvert}{2r+1}, r = 11$
P_4		$\frac{\sum_{i=2}^{N}\lvert intenzitás_i - intenzitás_{i-}}{Képkockákszáma}$
P_5		$\frac{\sum_{i=2}^{N}\lvert pitch_i - pitch_{i-1}\rvert}{Képkockákszáma}$
P_6		$\overline{keresztkorreláció_{pitch}}$
P_7		$\frac{P_1 + P_2 + P_5 + P_6 + P_8}{5}$
P_8		$P_4 * 200$
P_9		$var\big(FFT(ampl)\big)$
P_{10}		$var\big(FFT(pitch)\big)$
P_{11}		$\frac{P_7}{P_8}$
P_{12}		$\frac{P_3 * 100}{P_4}$
P_{13}		$\frac{P_{12}}{P_4}$
…		…
P_{17}		var(ampl)
P_{18}		std(ampl)
P_{19}		var(intenzitás)
P_{20}		std(intenzitás)
P_{21}		$\frac{P_5}{P_{19}} * 1000$
P_{22}		$\frac{P_1}{P_2}$
P_{23}		$\frac{P_4}{P_3}$
P_{24}		$\frac{P_4}{P_5}$
P_{25}		$\frac{P_{11}}{P_2}$
P_{26}		$\frac{P_{26}}{P_{18}}$
P_{27}		$\frac{P_{19} * 10}{P_{22}}$
P_{28}		$\frac{P_{19} * 10}{P_{23}}$
P_{29}		$P_{28} - P_{29,}$

P_{30}		$\frac{P_{28}}{P_{29}}$
P_{31}		$(P_{19}/P_{22})/P_1/P\ 2$
P_{32}		$P_{33} = \frac{P_{23}}{P_{25} * 100}$