

Szegedi Tudományegyetem
Földtudományok Doktori Iskola

Doktori (PhD) értekezés

**Hiperspektrális távérzékelési módszerek alkalmazhatósága
tájvizsgálatokban**

Csendes Bálint

Témavezetők:

Dr. Merényi Erzsébet, egyetemi tanár, Rice University

Dr. habil. Mucsi László, egyetemi docens, SZTE

Természeti Földrajzi és Geoinformatikai Tanszék

Szeged

2016

Tartalomjegyzék

1. Bevezetés	4
1.1 Problémafelvetés	5
1.2 A doktori kutatás menete	7
2. Irodalmi áttekintés.....	9
2.1 Hiperspektrális távérzékelés alkalmazása tájvizsgálatokban	9
2.2 Sávkiválasztás és adatredukció.....	11
2.3 Machine learning és SOM alkalmazások	12
2.4 Hazai hiperspektrális kutatások	13
2.5 Hiperspektrális tájvizsgálatok elméleti háttere.....	15
2.5.1 A hiperspektrális vegetáció vizsgálatok elméleti alapjai	15
2.5.2 A talajok spektrális jellegzetességei	16
2.5.3 Mezőgazdasági területek vizsgálatának sajátosságai	18
2.6 A hiperspektrális kutatások kiterjesztési lehetőségei.....	18
2.6.1 Hiperspektrális műholdfelvételek.....	19
3. Mintaterület, adatok, és a képfeldolgozási módszerek alapjai.....	21
3.1 A vizsgált mintaterület természetföldrajzi, tájtörténeti jellemzői	21
3.2 Adatok	24
3.2.1 Hiperspektrális adatok	24
3.2.2 Adatfeldolgozási környezet	26
3.3 Képfeldolgozási módszerek.....	27
3.3.1 A hiperspektrális felvételezés alapjai.....	27
3.3.2 A tanítóterületek kezelése	28
4. Eredmények	30
4.1 Az adatok előfeldolgozásának eredményei	30
4.1.1 Zajszűrés.....	30
4.1.2 Adatnormalizáció.....	32
4.2 Képi statisztikai vizsgálatok és adatredukciós eljárások	32
4.2.1 A vizsgált osztályok.....	32
4.2.2 A Hughes-hatás.....	33
4.2.3 Az adatredukció módjai	35

4.2.4 Főkomponens analízis	35
4.2.5 Spektrális szeparabilitási számítások.....	38
4.2.6 Spektrális diszkrimináció sávonként	41
4.2.7 A spektrális homogenitás vizsgálata	44
4.3 Irányított képosztályozások összehasonlítása	47
4.3.1 Az osztályozások előkészületei	47
4.3.2 A terepi vegetáció felmérés és ökológiai háttere	48
4.3.3 Mezőgazdasági adatgyűjtés	51
4.3.4 Felszínborítás térképezése spektrális szögeltérési osztályozással (SAM).....	52
4.3.5 Paraméteres osztályozási módszerek.....	56
4.3.6 Térképezés Maximum likelihood módszerrel	57
4.3.7 Képosztályozás Support Vector Machine alkalmazásával.....	60
4.3.8 SOM-hibrid osztályozás és klaszterezés.....	63
4.3.9 A felszínborítási térképezések diszkussziója	69
4.3.10 Mesterséges felszínek térképezése	71
4.3.11 Átmozgatott talajok osztályozása	72
4.3.12 A lombkorona alatti cserjeszint spektrális vizsgálata	73
4.4 Belvítérképezés irányítatlan és irányított képosztályozásokkal	76
4.4.1 A belvíz, mint környezeti probléma.....	76
4.4.2 Terepi belvízfelmérés.....	77
4.4.3 A távérzékeléses belvízkutatások előzményei	78
4.4.4 Hiperspektrális adataim a belvizes területekről.....	79
4.4.5 Irányított belvízosztályozások	80
4.4.6 Irányítatlan belvízosztályozás ISODATA módszerrel	84
4.4.7 Osztályozási eredmények összevetése a LiDAR domborzatmodellel.....	88
4.5 A statisztikai módszerek alkalmazása városi mintaterületeken	91
4.5.1 Statisztikai vizsgálatok városi mintaterületen	91
4.5.2 Kertvárosi térképezés laboratóriumi spektrumok alapján	94
5. Összefoglalás.....	97
Irodalomjegyzék	101
Köszönetnyilvánítás.....	114
Summary	115
Ábrák és táblázatok jegyzéke	119
Mellékletek	122

1. BEVEZETÉS

A földtudományi és környezeti vizsgálatok adatszükségletei újabb és újabb technológiai fejlesztésekre, innovációra ösztönzik a kutatóintézeteket és a mérőműszerek tervezőit, és bár napjainkban a rendelkezésre álló információ mennyisége exponenciálisan bővül, az igények a közeljövőben várhatóan tovább fognak növekedni olyan környezeti jelenségek megfigyelésében, melyeket a tudomány eddig eszközök, anyagi források, vagy érdeklődés hiányában nem tanulmányozott kellő mélységben (Haggett 2001).

Míg az elmúlt évtizedekben a földrajzi adatgyűjtés technikai feltételei folyamatosan javultak, egyre több és jobb minőségű terepi és távérzékelte adatot képesek a kutatók előállítani, addig az adatok interpretációjában, koherens környezeti modellek megalkotásában a sikeres kísérletek (környezeti kockázatbecslés, változás-vizsgálatok) ellenére még mindig jelentős a lemaradás (Skidmore 2002).

A földrajzi kutatások jellemzően a másodlagos források széles köréből merítik adataikat, az elérhető felvételek, terepen, illetve laboratóriumban mért adatsorok azonban gyakran nem illeszkednek teljesen a vizsgálatok céljaihoz, rosszabb esetben pedig a célkitűzéseket igazítják a rendelkezésre álló anyagokhoz (Montello & Sutton 2013).

Egyik legbőségesebb adatforrásunkat a távérzékelte felvételek jelentik, mivel feldolgozásuk egyszerű, időtakarékos és nagyrészt automatizált, továbbá kiértékelésük sem feltétlenül igényel mélyebb ismereteket, illetve laboratóriumi vizsgálatokkal kiegészítette terepi felméréseket az elemzett földrajzi jelenség okairól és összefüggéseiről. Az említettekből következően a képi adatok különösen alkalmasak környezeti folyamatok térbeli kiterjedésének becslésére is.

A távérzékelési technika jelentős fejlődésen ment át, a kezdetben hőlégballonokra, galambokra rögzítette képalkotó berendezéseket a 21. század elejére műholdplatformok és távirányítású rendszerek váltották fel és a képek felbontása is nagy mértékben finomodott, lehetővé téve a felszíni anyagok rendszeres, egyes fizikai-kémiai paraméterek kiértékelését is magában foglaló vizsgálatát.

A hiperspektrális, illetve LiDAR felvételezések elterjedésével megnövekedett adatmennyiség feldolgozása, értelmezése olyan kihívás elé állította a kutatókat, melynek

megoldására mind a mai napig nem sikerült minden hatékonysági és minőségi szempontot kielégítő módszereket kidolgozni (Merényi et al. 2007a, Merényi et al. 2014). Jelen doktori disszertáció tárgya a nagy spektrális felbontású adatok feldolgozásának módszertani áttekintése és továbbfejlesztése volt, azzal a céllal, hogy Magyarországra jellemző környezeti problémák, tájváltozásokkal összefüggő földrajzi jelenségek távérzékeléses elemzéséhez minél pontosabb és hatékonyabb eljárásokat fejlesszek ki.

1.1 Problémafelvetés

Távérzékeléses vizsgálatokat szinte már az első légi járművek megjelenése óta használnak a földrajzi kutatásokban, az utóbbi évtizedekben viszont, köszönhetően az olyan új fejlesztéseknek, mint a CCD chip, a költséghatékony műholdplatformok, a felvételezések közben már rendelkezésre álló GPS adatok és a korábbiaknál nagyságrendekkel bővebb számítási kapacitás, új szintre lépett a légi és űrfelvételezési technológia. A nagyfelbontású hiperspektrális képeket már széleskörűen alkalmazzák földrajzi vizsgálatokban, a hatalmas, négyzetkilométerre vetítve a repülési magasságtól függően akár több gigabyte-ot kitevő adatmennyiség sokoldalú elemzése egyúttal megköveteli a statisztikai módszertan újragondolását is. A multispektrális felvételekhez használt megközelítések mindehhez már nem feltétlenül célravezetők, az esetek többségében pedig, ha kellő számú mintát sikerül azonosítani, összetett indexszámítások nélkül is megbízhatóan elkülöníthetők a vizsgált felszíntípusok.

Dolgozatomban a Szeged mellett található Tápai-rét példáján keresztül kívántam bemutatni a több száz spektrális sáv feldolgozására és interpretálására alkalmas képfeldolgozási, illetve osztályozó algoritmusokat, különös tekintettel azok módszertani korlátaira, melyek jórészt a gépi tanulási módszerek túlilleszkedéséből (overfitting) fakadnak. A már létező módszerek magyarországi alkalmazása mellett olyan újfajta statisztikai eljárások kifejlesztését jelöltem ki feladatommak, melyek növelik a spektrális diszkrimináció hatékonyságát és bővebb képi információt nyújtanak a térképezett felszíni anyagokról akkor is, ha a tanítópontok száma erősen korlátozott.

Célkitűzéseim az alábbiak voltak:

- 1, *2010 szeptemberében készített AISA hiperspektrális felvételeimet tájvizsgálatokra kívántam használni, amire a Szeged melletti Tápai-rét sajátos környezeti problémái kapcsán nyílt lehetőségem. A Tisza és a Maros torkolata fölött található mintaterületen több, földrajzi szempontból releváns tájváltozási jelenség is kutatható a nagyfelbontású adatokkal, így az özönnövények terjedése, a belvízképződés, valamint a szénhidrogén-bányászattal járó antropogén felszínformálás.*
- 2, *A különböző felszínborítási kategóriához tartozó spektrumgörbék elkülönítésére számos eljárás ismert a nemzetközi tudományos irodalomban. Céлом az volt, hogy megvizsgáljam és összehasonlítsam a spektrális diszkrimináció módszereit és a megfelelő adattranszformációk alkalmazásával csökkentsem a felvételek zajtartalmát.*
- 3, *Doktori kutatásom további céljának azt jelöltem ki, hogy az adatminták spektrális homogenitásának vizsgálatára újfajta módszert fejlesszek ki, ami a több száz spektrális sávban felvételezett hiperspektrális képekre is alkalmazható. Az egy osztályhoz tartozó képelemek reflektancia értékeinek szórása meghatározza az osztályozás pontosságát, a homogenitást leíró statisztikai adatok egyúttal fontos információt nyújthatnak a vizsgált anyagokról.*
- 4, *A kijelölt tanítóterületek ellenőrzésére a távérzékeléses kutatásokban gyakran használnak úgynevezett spektrális szeparabilitási számításokat. A hiperspektrális adataimon terepi felméréseket követően felvett mintapontokat Transformed Divergence és Jeffries-Matusita-féle szeparabilitási indexekkel kívántam kiértékelni.*
- 5, *Feladatomban jelöltem ki, hogy többféle, paraméteres és nem-paraméteres képosztályozási algoritmust is összehasonlítsak mind elméleti hátterük, mind alkalmazhatóságuk szempontjából. Mintaterületem, a Tápai-rét távérzékeléses*

térképezésében komoly kihívást jelent a vegetációs felszínnek mozaikos szerkezete és az egyes foltok viszonylag kis kiterjedése, ezért különös hangsúlyt kell fektetni a korlátozott számú mintával is használható eljárások alkalmazására. A pontosságot kappa-indexszel és konfúziós mátrixszal terveztem számszerűsíteni.

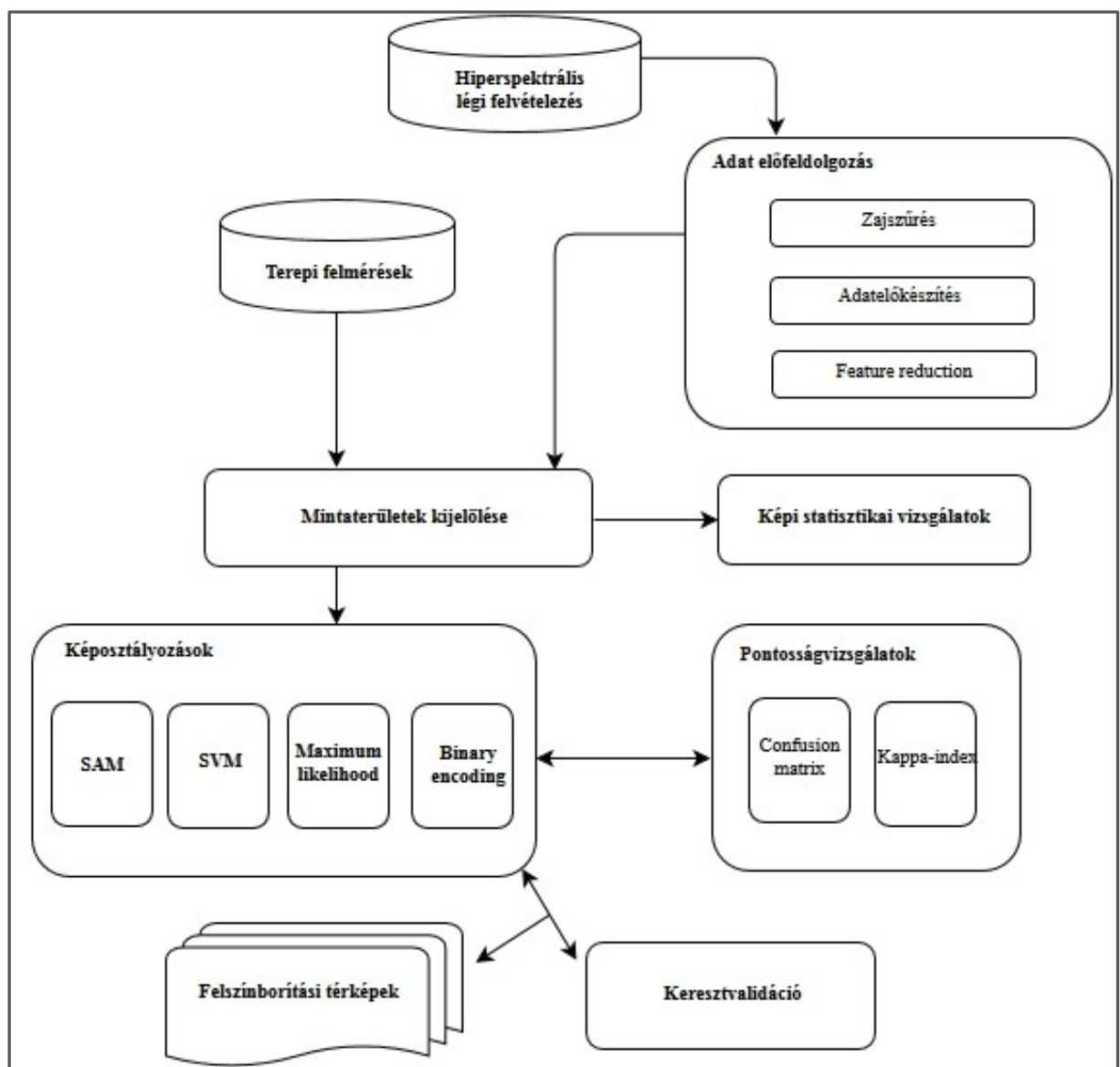
6, *Célkitűzéseim között szerepelt továbbá, hogy hiperspektrális adataimon megvizsgáljam az irányítatlan és az irányított osztályozások lehetőségeit és korlátait. A belvízzel, illetve mezőgazdasági és természetes növényfajokkal fedett felszíneket választottam mintaterületnek, figyelembe véve a térképezett osztályok térbeli mintázatának sajátosságait és az esetleges képi zajból adódó spektrális anomáliákat. Választ kívántam adni arra a kérdésre, hogy az egyes távérzékeléses tájvizsgálati feladatokra az osztályozási pontosság és a terepi adatgyűjtés szempontjából melyik megközelítést érdemes alkalmazni.*

1.2 A doktori kutatás menete

A kutatáshoz szükséges adatokat három forrásból merítettem, a távérzékelte adatokat terepi növény- és egyéb tájvizsgálatokkal egészítettem ki, valamint amikor arra lehetőségem nyílt, a vizsgált felszíntípusok mintáinak laboratóriumban mért spektrális adatait is felhasználtam.

A képi adatok előkészítéséhez zajszűrést alkalmaztam, eltávolítottam a hibás adatsávokat, illetve adat-transzformációval mérsékeltem a redundanciát. Az előfeldolgozási lépések egyszerre szolgálták a pontosság javítását és a szükséges számítási kapacitás csökkentését. Tanítóterületeimen, melyeket a terepi felmérések adatai alapján jelöltem ki, különféle statisztikai vizsgálatok végeztem, mindenekelőtt az egy osztályhoz tartozó képelemek spektrális homogenitását kívántam megbecsülni az eredményesebb klasszifikáció érdekében. Önmagukban is fontos információval szolgálnak a spektrális statisztikai mutatók, a reflektancia-görbék sajátosságain túl a vizsgált anyagok fizikai, kémiai tulajdonságairól is. A tanító-ellenőrző pontok szétosztásának lehetséges térbeli aránytalanságait keresztvalidációval küszöböltem ki, így képosztályozásaimat egy általam kialakított, valamint tovább öt, véletlenszerűen generált pontkiosztási halmazon végeztem, eredményeiket pedig tévesztési

mátrixszal, valamint Cohen-féle kappa indexszel értékeltem ki. A pontossági adatokat és az elkészült felszínborítási térképeket a képi statisztikák mutatószámáival vetettem össze és módszertani következtetéseimet is ezek alapján vontam le. Az eredménytérképek vizuális kiértékelése is fontos információval szolgált a vizsgált algoritmusokról, különösen a túlilleszkedés mértékét lehetett a félreosztályozott területeken nyomonkövetni. Az 1. ábrán látható sémát követtem a Szeged városi, a tápai-réti és a belvízzel elöntött mintaterületek térképezésénél is. Dolgozatomban a felhasznált módszerek magyar megnevezését részesítettem előnyben, általánosan elterjedt kifejezés hiányában azonban az angolul hivatkoztam ezekre.



1. ábra: A doktori kutatás adatforrásai és fontosabb lépései

2. IRODALMI ÁTTEKINTÉS

2.1 Hiperspektrális távérzékelés alkalmazása tájvizsgálatokban

A hiperspektrális technika terjedésével, az olcsóbb és hatékonyabb felvételező berendezések, valamint a műhold platformra szerelt szenzorok megjelenésével folyamatosan bővül az alkalmazási területek köre is. A következőkben kivonatolt kívántam adni a kurrens hiperspektrális térképezési kutatásokról, bemutatva a konkrét földrajzi, illetve környezeti problémákat, kérdésfelvetéseket és az elemzések általánosabb érvényű következtetéseit.

A Yellowstone Nemzeti Park területén végzett felmérés eredményeit foglalja össze *Fairweather et al. (2012)*, összevetve a multi- és hiperspektrális felvételek osztályozási eredményeit. A szemi-arid klímájú terület vegetációját jelentős éven belüli változékonyság jellemzi, ezért a kutatók számára fontos volt az alkalmas felvételezési időpontok megválasztása. A nagyobb időbeli felbontás hátulütőjeként a mozaikos szerkezetű növényfelszínek térképezését alacsonyabb geometriai felbontású adatokon kellett elvégezni. Spektrális keveredés nem csupán az egyes növényfajok között jelentkezett, hanem a vizsgált *Artemisia* cserjék nem fotoszintetizáló részei és a fotoszintetikus, arid környezetre jellemző vegetáció, valamint a világosabb talajfelszínek között is. Ugyanakkor a szerzők eredményei alapján önmagában már az AVIRIS felvételek finomabb spektrális felbontása is lehetővé teszi a kevert képelemek nagy pontosságú osztályozását.

A hiperspektrális növényterképezés egyik, nem teljesen megoldott problémája a fenológiai fázisok azonosítása. A lipcsei környezetkutatási központ (Helmholtz-Zentrum für Umweltforschung) munkatársai azonban laboratóriumi körülmények között sikeresen különítették el a BBCH skála bizonyos stádiumait különféle biokémiai és biofizikai paraméterek alapján, az árpa példáján. Eredményeik szerint az elterjedtebb mutatók közül a víztartalom (GWC) a legalkalmasabb az adatok osztályozására, a módszert pedig a terepi, távérzékeléses felmérésekre is ki lehet terjeszteni (Lausch et al. 2015).

Santos et al. (2010) élőhely-terképezés céljából vizsgálta a georgiai fenyvesek állapotát. Miután ANOVA statisztikával sávonként értékelték ki az előregedő, a kiszáradt és

az egyéb állományok szeparabilitását, CART osztályozási fákkal (Classification and Regression Tree) különítették el azokat a felvételeken. Szintén regressziós számításokkal értékelték ki az erdőtűz után telepített facsemeték állományának növekedését, a Yellowstone Nemzeti Park leégett erdeinek regenerálódását (Potter et al. 2012). Széles irodalma van a kifejezetten invazív növények terjedésével és felmérésével foglalkozó tudományos közleményeknek, trópusi területeken (Joshi et al. 2005) és a mérsékelt övben is (Hestir et al. 2008, Andrew & Ustin 2008).

A szervesanyag tartalom becslésében sikereket ért el *Santini et al. (2010)* is, amikor Velence környéki lagúnák klorofill, tripton és a színes oldott szervesanyagok (CDOM) koncentrációját becsülték meg Hyperion, CASI és MIVIS felvételeken, egy, mindenekelőtt az abszorpciós sávokat vizsgáló bio-optikai modell alapján, majd a modell helyességét laboratóriumi in situ vízmintákkal is verifikálták. Szintén sikeres kutatások folynak a növényi tápanyag tartalom becslésében, melyeknél a hiperspektrális adatokat elsősorban neurális hálózattal dolgozzák fel (Knox et al. 2011, Mutanga & Skidmore 2004, Mutanga & Kumar 2007).

A nemzetközi irodalomban találhatunk példát a hiperspektrális távérzékelés oceanográfiai alkalmazásaira is, mint Torrecilla és társai kutatásában, ahol a tengeri fitoplankton optikai tulajdonságait vizsgálták (Torrecilla et al. 2011). Szintén hajó fedélzetéről készített felvételeket automatizált reflektancia méréseihez *Simis & Olsson (2013)*.

A hiperspektrális technika alkalmazható más égitestek vizsgálatára is, így tudományos igényű felvételek készültek a Holdról, a Szaturnuszról és a Marsról is (Chauhan et al. 2015, Merényi et al. 2004). Az így nyert képek értelmezéséhez olyan, földi referenciaterületet érdemes tanulmányozni, melynek litológiai, hőmérsékleti és egyéb környezeti jellemzői hasonlítanak a vizsgált bolygóéhoz. *Gleeson et al. (2010)* a Galileo űrszonda NIMS spektrométerével a Jupiter holdjáról, az Európáról rögzített adatait vetette össze egy kanadai, sarkvidéki mintaterület kénben gazdag üledékeivel és számottevő hasonlóságot fedezett fel a hold egyes felszíni anyagainak spektrumgörbéivel.

2.2 Sávkiválasztás és adatredukció

Hyperion űrfelvételek kevert pixeleit vizsgálta *Somers & Asner (2013)*, a találati pontosság növelésére pedig egy automatizált eljárással, szeparabilitásuk, ill. korrelációjuk alapján állították össze a spektrális szétválasztásba bevont csatornák számát, minek köszönhetően a számítási idő is érdemben csökkent. A megfelelő sávok kiválasztását, mint az adatredukció (feature reduction) egyik lehetséges módszerét *Inoue et al. (2012)* a növényi nitrogéntartalom (CNC) térképezésére is sikeresen használta.

A redundáns képi információ eltávolításának másik módja a főkomponens analízis, vagy annak továbbfejlesztett változata, az MNF (Minimum Noise Fraction) transzformáció. *Zhang & Xie (2012)* következtetései szerint az MNF-fel zajmentesített sávok, amellet, hogy csökkentik a vizsgálatok számítási igényeit, a hiperspektrális távérzékelésen alapuló vegetációvizsgálatokban is különösen hasznosak, ahogyan ezt már korábbi tanulmányok is alátámasztották (Belluco et al. 2006, Yang et al. 2009).

A szeparabilitási számítások célja távérzékelési alkalmazásokban az osztályok elkülöníthetőségének számszerűsítése, gyakran ANOVA és T-próba vizsgálatok kiegészítésére használják, a kétfajta módszer kombinációjával átfogó képet kaphatunk az osztályokon belüli és az osztályok közötti varianciáról, több sáv esetén is (Vaiphasa et al. 2005, Andrew & Ustin 2008, Ustin et al. 2009, Vaiphasa et al. 2007).

A területi autokorreláció alapján kiválasztott spektrális sávok, sávkombinációk szeparabilitását vizsgálta a Jeffries-Matusita módszerrel *Warner & Shank (1997)*, megfigyeléseik szerint egyes hiperspektrális csatornák hozzáadása nehezítheti is az osztályok elkülönítését.

Optimális sávkombinációkat értékelt ki *Fernandes et al. (2013)* is, a CART eljárással kiválogatott tartományokon Bhattacharya- és Jeffries-Matusita-féle számításokat is végzett, hogy elkülönítse a vizsgált olasz nád állományokat a többi növényfajtól, terepi spektrométeres adatokon.

2.3 Machine learning és SOM alkalmazások

Komplex, magas dimenzionalitású adatstruktúra feldolgozásához jellemzően a machine learning alkalmazásokat részesítik előnyben a távérzékelés területén is (Lary et al. 2015, Merényi et al. 2007a, Merényi et al. 2009). Elméleti tanulmányában *Lary et al. (2015)* a gépi tanulási módszerek előnyeit mutatták be „big data” alkalmazásokban, következtetéseiket aeroszolk forrásainak távérzékeléses térképezésével támasztották alá.

A neurális hálózatok sajátossága az adatok nemlineáris transzformációja, aminek köszönhetően pontosabban térképezhetők a hiperspektrális felvételek osztályai (Licciardi et al. 2012, Benediktsson et al. 1994) míg a sávselekción és a főkomponens analízis szignifikáns információ veszteséghez vezethet (Howell et al. 1994, Moon & Merényi 1995, Merényi et al. 2014).

A többrétegű perceptronon és az egyszerűbb backpropagation algoritmusokon túl számos további ANN-alapú osztályozási módszert ismer a képfeldolgozás irodalma. Adaptív, gépi tanulási módszerrel kísérletezett *Zhong & Zhang (2012)*, amikor hiper- és multispektrális adatokon sikeresen alkalmazták az immunrendszer antitestjeinek működését modellező ABNet rendszert.

Liu et al. (2010b) laboratóriumi spektrumok deriváltjaiból előállított főkomponens sávokon learning vector quantization-nal (LVQ) térképezte a rizsvirágzat gombafertőzéseit. Szintén adaptív LVQ osztályozást választott *Zhang & Xie (2012)* az Everglades mocsárvidék vegetáció felméréseihez, következtetéseikben a nagyfelbontású adatok térbeli textúrájának fontosságát hangsúlyozták.

A Self-organizing Map (SOM) osztályozás előnyei különösen azoknál az alkalmazásoknál mutatkoznak meg, ahol a magas komplexitású adat belső összefüggéseit előzetesen klaszterezéssel kell feltárni, illetve ahol a vizsgált anyagok fizikai-kémiai tulajdonságai ezt megkövetelik, mint egyes geokémiai kutatásoknál (Penn 2005).

Merényi et al. (2007a) Ocean city városi területeiről készített Daedalus multispektrális és AVIRIS hiperspektrális felvételeken végzett vizsgálataik kapcsán állapították meg, hogy magasabb adatkomplexitásnál hangsúlyosabbak az SOM megközelítés előnyei (szemben a konvencionális klaszterezésekkel, mint az ISODATA), eredményének feldolgozása viszont

részben manuálisan történik, ezért hatékony kiértékelési vizsgálatokra van szükség, mint amilyen az általuk kifejlesztett ConnVIS.

A topológiai térképezés jelentőségére hívja fel a figyelmet *Villmann et al. (2003)* a magas dimenzionalitású adatoknál, a továbbfejlesztett SOM funkciókat (magnification control, generalized relevance learning vector quantization) mindenekelőtt a pontosság növelésére használták. A hibrid osztályozás ugyanakkor zajos adaton, nagy számú osztály mellett is megbízható eredményt ad, esetenként pontosabbat, mint a statisztikai megközelítésű algoritmusok, az SVM, a Mahalobis Distance, Maximum likelihood vagy a minimum Euclidean distance (Merényi et al. 2014).

A térképezési feladat és a felhasznált adat sajátosságaihoz igazodva további adatfeldolgozási módszerekkel lehet bővíteni az SOM vizsgálat eszköztárát, a spektrális szétválasztáshoz érdemes lehet a hibrid osztályozást fuzzy klaszterezéssel kombinálni (Liu et al. 2010a), míg *Moshou et al. (2001)* local linear mapping (LLM) algoritmussal pontosította a nehezebben elkülöníthető mezőgazdasági gyomok azonosítását.

Űrmissziókhöz fejlesztette ki a HyperEye adatfeldolgozási rendszert *Merényi et al. (2007b)*, céljuk a földi irányítástól távol zajló műveletek lehető legmagasabb szintű automatizációja volt. Az osztályok elkülönítésének hatékonyságában jelentős előrelépést értek el, egyes munkafázisokhoz, mint a klaszterek és a vizsgált anyagok megfeleltetéséhez azonban továbbra is operátorok szükségesek (Merényi et al. 2007b).

2.4 Hazai hiperspektrális kutatások

A legelső, Magyarország területén végzett spektrométeres térképezésre 1989-ben került sor a kanadai MONITEQ műszereivel a Szigetköz fölött, kapcsolódva a Bős-nagymarosi vízlépcső beruházáshoz (Büttner 1990). A HYSSENS 2002 volt az első magyar hiperspektrális távérzékelési projekt, melyet a Magyar Állami Földtani Intézet szervezett meg Kardeván Péter szakmai irányításával (Kardeván et al. 2003). A vizsgálatok elsősorban a hazai környezeti veszélyforrások felmérését szolgálták, a program keretein belül légi képek készültek a Mátra egykori bányaterületeiről, a Sajó alsó, nehézipari eredetű környezeti terhelésnek, szennyezésnek kitett árteréről, valamint Debrecen közelében található,

mezőgazdasági telepekről. A program eredményeit és tanulságait Hargitai Henrik foglalta össze (Hargitai 2006).

Az ELTE TTK Természetföldrajzi Tanszékén a multispektrális adatok mellett műholdról készített hiperspektrális felvételek (Hyperion EO-1) elemzésével is foglalkoznak. Deák Márton és Mari László elsősorban a mesterséges felszínek, települések azonosításában ért el eredményeket (Deák et al. 2012), az alkalmazott spektrális szétválasztási módszereket pedig a növényzet osztályozására is felhasználták (Deák et al. 2014).

DAIS felvételeken vizsgálták Gyöngyös városi növényfelszíneit Jung András és társai. A képosztályozások mellett a hiperspektrális sávokra szabott vegetációs indexeket is alkalmazták, köztük a hNDVI-t és vegetációs stressz analízist (Jung et al. 2005).

Elsősorban a hiperspektrális képfeldolgozás elméleti kérdéseiben és osztályozási algoritmusok fejlesztéséről közölt eredményeket Berke József és Kozma-Bognár Veronika. A Berke nevéhez köthető Spectral Fractal Dimension (spektrális fraktáldimenzió, SFD) algoritmus, köszönhetően a logaritmussal számított értékeinek, kevésbé érzékeny a légköri, domborzati és egyéb hatásokra, anomáliákra (Kozma-Bognár & Berke 2010), az újfajta megközelítést pedig a zajszűrésben, illetve az optimális sávok kiválasztásában is hasznosítani tudták (Kozma-Bognár & Berke 2013).

A legjelentősebb hazai, hiperspektrális távérzékelési felmérésekkel foglalkozó műhely Gyöngyösön, a Károly Róbert Főiskolán működik. A kutatócsoport első eredményei mezőgazdasági területek térképezéséhez köthetők (Burai 2006), AISA felvételező műszerükkel környezeti vizsgálatokat végeztek ártéri, hegyvidéki és szikes élőhelyeken is, melyek célja elsősorban a növényzet egészségi állapotának felmérése, illetve az özőnfajok térképezése volt (Burai & Tomor 2011, Burai et al. 2014, Kertész et al. 2014). Városi kutatásukban Sopron példáján sikerült szignifikáns kapcsolatot ($R^2=0.6288$, $P<0.05$) kimutatniuk a légi és terepi spektrométerekkel felvételezett, first derivative spektrális indexek és a laboratóriumban mért klorofill tartalom között (Visi-Rajczi et al. 2012). A legpontosabb osztályozásokat SVM módszerrel állították elő (Kertész et al. 2014, Burai et al. 2014), egyes alkalmazásoknál ugyanakkor a szögeltérési térképezés (SAM) is használhatónak bizonyult (Burai et al. 2010).

A szeparabilitási számításokat hazánkban elsősorban a tanítóterületek kiértékelésére alkalmazzák (Tobak 2013), egyik első felhasználása multispektrális SPOT adatokon végzett felszínértékeléshez kötődik (Büttner et al. 1988).

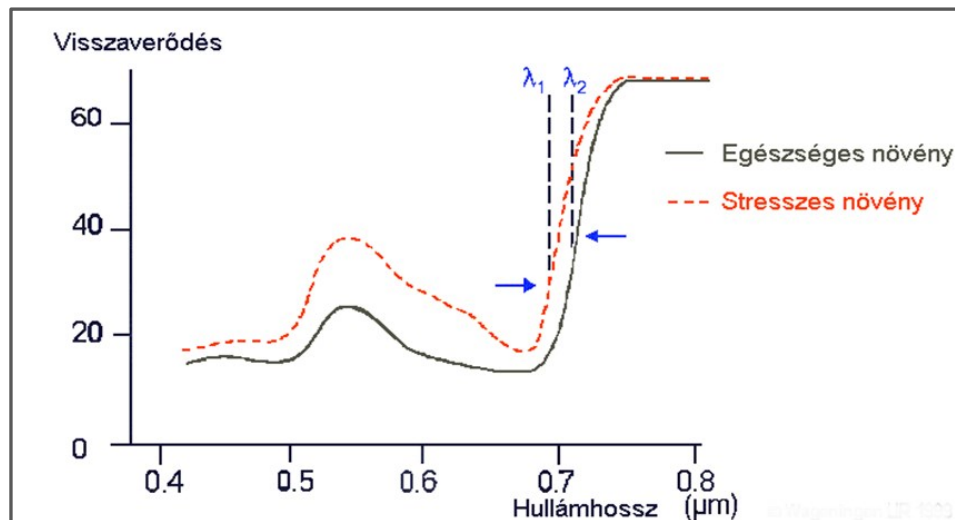
2.5 A hiperspektrális tájvizsgálatok elméleti háttere

2.5.1 A hiperspektrális vegetáció vizsgálatok elméleti alapjai

A FAO 2010 évi erdőfelmérése (Global Forest Resources Assessment) szerint a földfelszín 31 százalékát borítják erdők és további 9 százalékát olyan fás területek, ahol a lombkoronával fedett foltok aránya meghaladja a 10 %-ot, vagy 5 százalék feletti érték mellett a fák lehetséges magassága eléri az 5 métert. A természetes vegetáció összetétele és állapota, mint fontos környezeti indikátor, az egyik legfontosabb témája a távérzékeléses kutatásoknak. A növényzetről gyűjtött spektrumgörbéken már ritkábban felfedezhetők fel olyan markáns abszorpciós sávok, mint az ásványoknál, ezért térképezésüknél a spektrális jellegzetességek azonosítása helyett egyéb, parametrikus és nonparametrikus képi statisztikai megközelítések kerülnek előtérbe. Az egyes növényfajokat gyakran a terepi felmérésen is nehéz egymástól elkülöníteni, a spektrális diszkrimináció pontossága rendszerint alacsonyabb a földtani alkalmazásoknál megszokott értékeknél.

A vegetációs spektrumgörbék a látható tartományban mutatják a legnagyobb változatosságot, köszönhetően a klorofill, a xantofill, a karotin és az egyéb pigmentek abszorpciós tulajdonságainak, a közeli infravörös tartományban viszont már a beérkező energia mintegy fele visszaverődik (Clevers 1994). A két hullámhossz tartomány határán figyelhető meg az ún. „vörös él” (red edge), ami a visszaverődő fény arányának hirtelen megugrására utal a 680 és 750 nanométer között (Horler et al. 1983). A vörös él pozíciójára csak minimális befolyással vannak a talaj és az atmoszférikus hatások, ellenben a növényi stresszhatásokat jól jelzi eltolódása az alacsonyabb hullámhosszúságú sugárzás irányába (blue shift, 2. ábra), amit tíz nanométer alatti spektrális felbontású adaton lehet megbízhatóan azonosítani (Clevers 1994).

A hiperspektrális felvételek bővebb adattartalma egyértelműen pontosabb növényterképezést és levélfelületi becslést (LAI) tesz lehetővé, mint a szélessávú, multispektrális műholdképek (Lee et al. 2004), fennáll ugyanakkor a lehetősége annak, hogy a nagyszámú spektrális csatorna felhasználásával túlillesztetté válnak a modellek (Thenkabail et al. 2000).

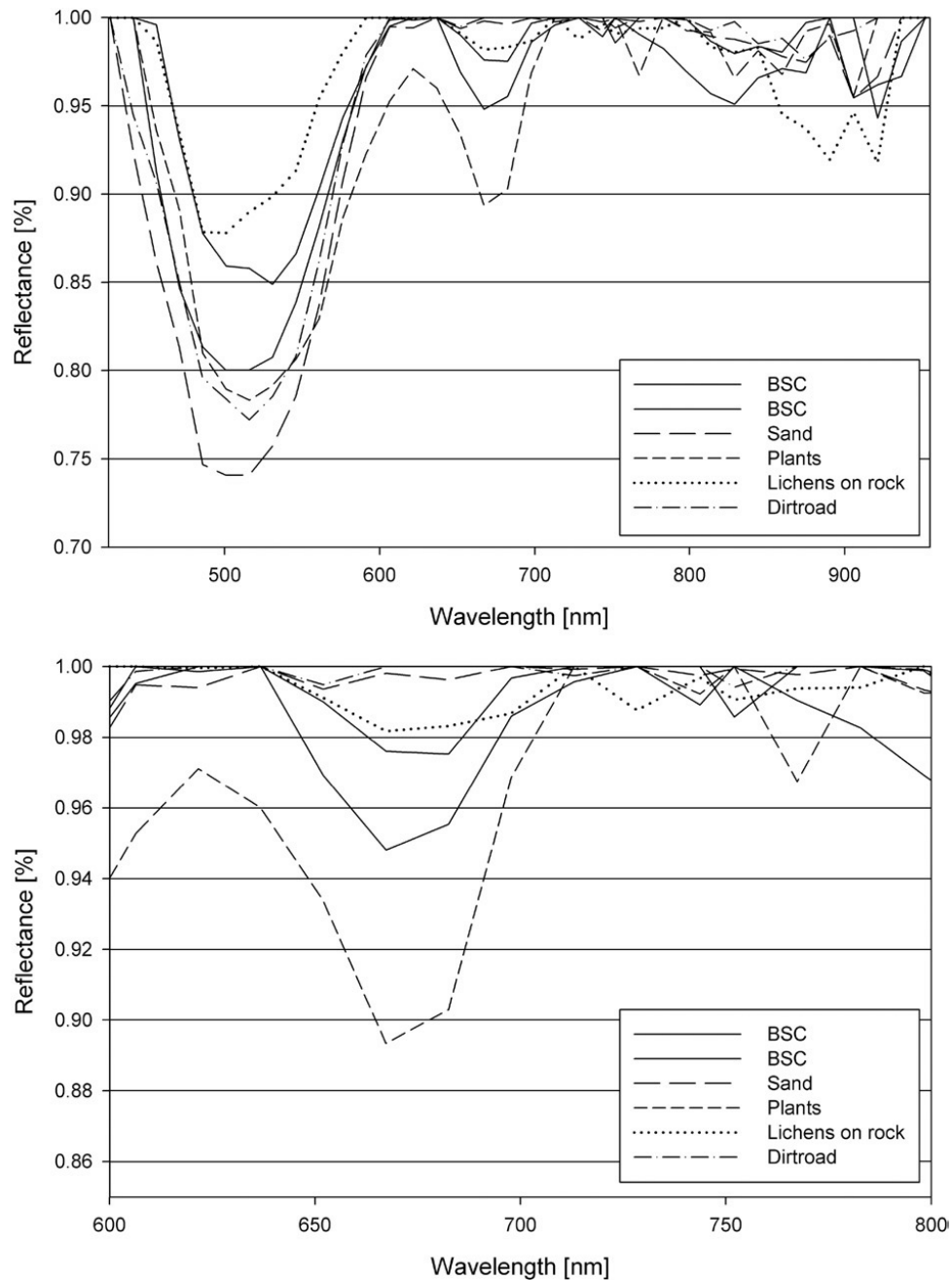


2. ábra: Egészséges és stressznek kitett vegetáció spektrumgörbéje, a reflektancia százalékában
(forrás: Burai 2007)

2.5.2 A talajok spektrális jellegzetességei

Tudományos szempontból különösen nagy jelentősége van a kriptobiotikus kéreg kutatásának. A BSC (Biological Soil Crust) néven is ismert legfelső, vékony talajréteget a benne található cianobaktériumok, gombák és egyéb mikroorganizmusok formálják, szervesanyag tartalma és a benne tárolt vízkészlet révén az ökoszisztéma fontos része, arid és szemi-arid klímájú területeken az élővilág szempontjából kiemelkedő jelentősége van (St. Clair & Johansen 1993). A növényzettel nem borított talajokon hiperspektrális technológiával jól térképezhetők a BSC-k, térbeli felbontástól függően spektrális szétválasztással (Stuart, 2012), vagy az abszorpciós sávok azonosításával (O'Neill 1994, Weber et al. 2008).

A 3. ábrán láthatók egy fűsivatagi éghajlatú mintaterület főbb felszíntípusainak abszorpciós sávjai egy hipotetikus, folytonos spektrumgörbéhez viszonyítva (continuum removed spectrum). A reflektanciaértékek változásaiban, anomáliáiban nyomon követhetők a talajkérget érő antropogén eredetű külső hatások, a taposás, a legeltetés és a gépjárművek okozta fizikai talajdegradáció (Johnston 1997, Stuart 2012).



3. ábra: A kriptobiotikus kéreg (BSC), a földutak és a szemiárid területek jellegzetes felszíntípusainak spektrumgörbéi, kontinuum-eltávolítás után (forrás: Weber et al. 2008)

2.5.3 Mezőgazdasági területek vizsgálatának sajátosságai

A szántóföldi növények felvételezése és térképezése némileg más megközelítést igényel, mint a természetes vegetációval fedett területek felmérése, a nagy kiterjedésű, homogén borítottságú mezőgazdasági parcellák azonosítása, terepi felmérése és az eredménytérképek kiértékelése viszonylag egyszerű, emellett új módszerek, osztályozási algoritmusok tesztelésére is kiválóan alkalmasak, mint a távérzékelés nemzetközi irodalmában sűrűn citált, AVIRIS szenzorral felmért Indian Pine mintaterület, vagy Magyarországon a tedeji gazdaság földjei (Kardeván et al. 2003).

Digitális légifotók és egyéb, földrajzi vonatkozású mérési adatok birtokában lehetőség van precíziós mezőgazdasági tervezésre (Precision Crop Management, PCM), melynek lényege a növényvédő szerek, a víz, műtrágyák és más komponensek optimalizált felhasználásában rejlik (variable rate technology), amihez alapvető szükség van megfelelő monitoring-eljárásokra (Robert 1997).

2.6 A hiperspektrális kutatások kiterjesztési lehetőségei

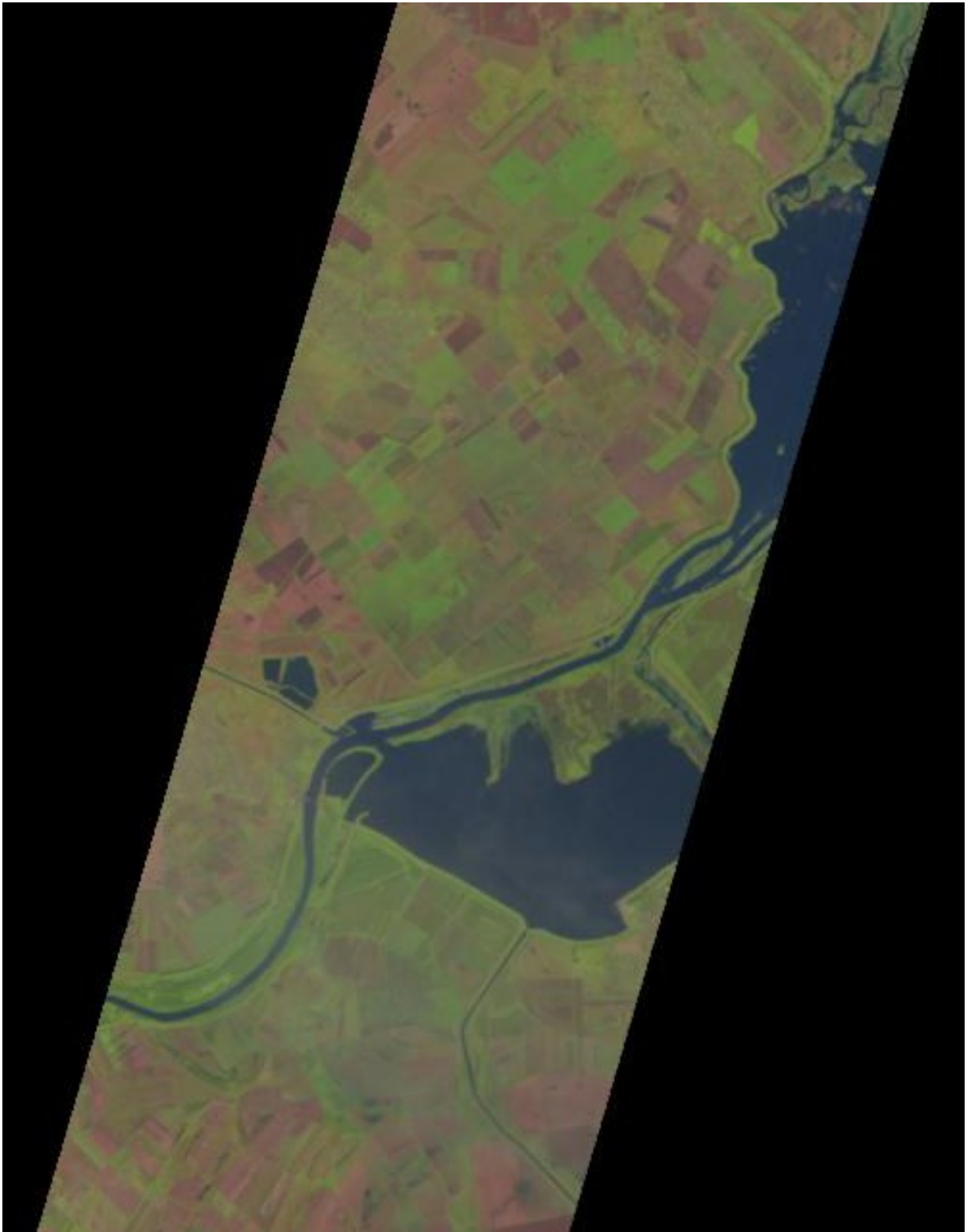
A nagyfelbontású adatok felvételezésénél gyakran merülnek fel olyan technikai korlátok, mint a szenzor érzékenységből vagy az alacsony repülési magasságból adódó memória, illetve térbeli megszorítások, melyek limitálják a környezeti elemzésekben hasznosítható adat mennyiségét.

A hiperspektrális felmérések kiegészítésére adatfúziós eljárásokat célszerű alkalmazni, ahol az adott esetben rosszabb minőségű képek magasabb geometriai vagy időfelbontásával pontosítható a térképezés (Zhang 2015). Távérzékeléssel nyert adatforrásainkon túl az anyagi összetétel laboratóriumi elemzésével, illetve in situ növénybiológiai vizsgálatokkal is nyerhetünk olyan információkat, amelyek már nagyfelbontású képeken is értelmezhetők (Lausch et al. 2016, Pause et al. 2016, Visi-Rajczi et al. 2012).

2.6.1 Hiperspektrális műholdfelvételek

Napjainkban már műholdról felvételezett hiperspektrális képek is hozzáférhetők a kutatók számára, az egyik legismertebb, világűrbe telepített képalkotó spektrométer a NASA EO-1 Hyperion szenzora 220 spektrális sávban, 30 méteres térbeli felbontással készít felvételeket, melyek egy része ingyenesen is elérhető az USGS (United States Geological Survey) adatbázisából.

A 4. ábrán a Kiskörei-víztároló nyugati fele látható, egy 2005-ös Hyperion felvétel kivágatán, ahol a hiányos zajszűrés ellenére tisztán felismerhetők a mezőgazdasági telekhatárok és a vízfelületek. Bár a nagyobb léptékű térképezések lényegesen több terepi felmérést igényelnek, az űrfelvételek segítségével nagy kiterjedésű, akár országrész méretű területeken lehetséges mind a mesterséges, mind a természetes felszínek megbízható osztályozása (Deák et al. 2012, Deák et al. 2014).



4. ábra: EO-1 Hyperion műholdfelvétel a Tisza-tó nyugati feléről (2005)

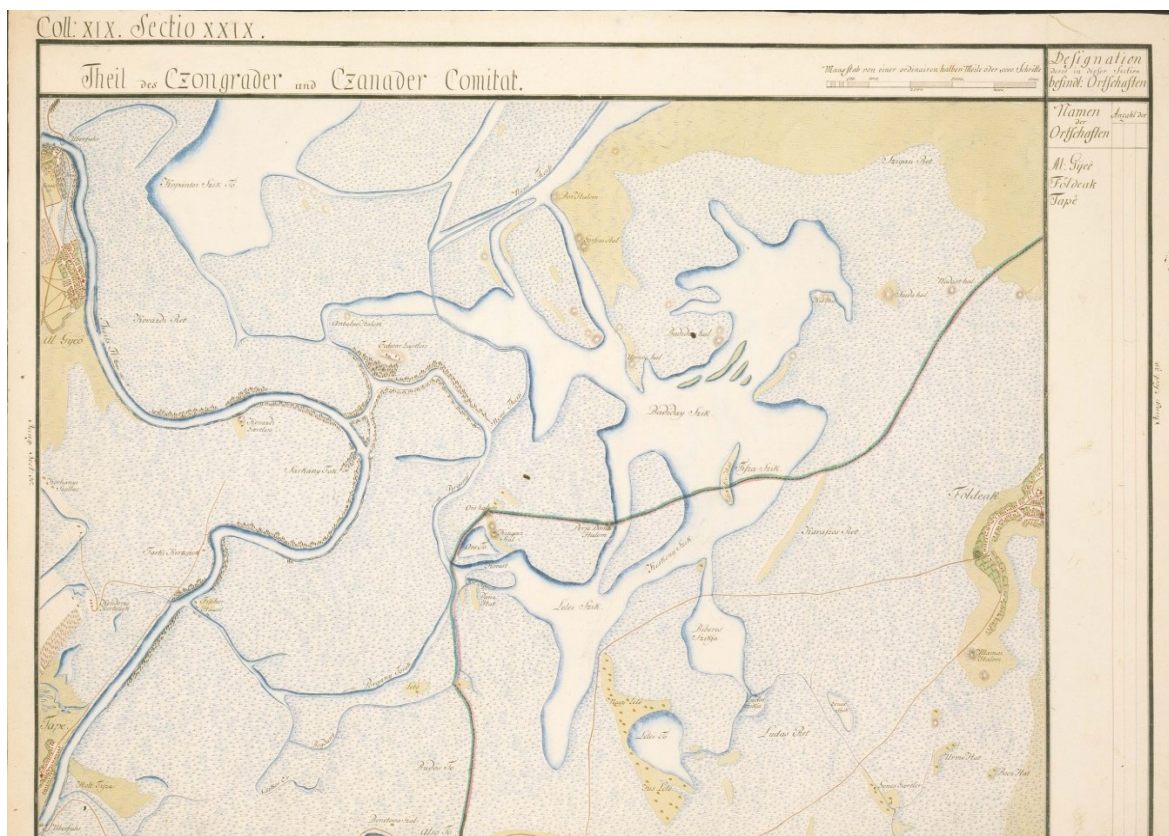
3. MINTATERÜLET, ADATOK, ÉS A KÉPFELDOLGOZÁSI MÓDSZEREK ALAPJAI

3.1 A vizsgált mintaterület természetföldrajzi, tájtörténeti jellemzői

Kutatásaimat Tápairéten, a földrajzi kataszteri rendszerben Marosszögnek nevezett kistáj (Dövényi 2010) nyugati részén végeztem (6. ábra). A terület közvetlenül a Tisza-Maros torkolattól északkeletre húzódik, északi határát a nagyfai Tisza-holtágnál, illetve a 2011-ben átadott M43-as autópályánál húzhatjuk meg, míg keleti szegélyét Vetyehát és az őskori leletekben gazdag Lebő-halom jelzi, melyen túl már Maroslele szántóföldjeit találjuk (7. ábra).

Ahogy neve is mutatja, Tápairét egykor Tápé falu kaszálója volt, elsősorban állattartás céljára hasznosították. Ennek oka, hogy a folyószabályozások előtt a területet időszakosan víz borította, ami lehetetlenné tette a tartós szántóföldi művelést, ugyanakkor gyümölcstermesztés is csak a magasártéren volt jellemző (ifj. Lele 1978). A 18. századi állapotok visszaköszönnek az I. katonai felmérés térképsorozatán is (5. ábra). A Maros ebben az időben még mai medrénél délebbre futott, a Tápairét jelentős részét pedig időszakos és állandó vizek valamint kaszálórétek foglalták el, melyek szintjéből csak néhány halom emelkedett ki.

A táj földtani viszonyait az alulról felfele finomodó holocén folyóvízi üledékek határozzák meg, melyeket csak ritkán szakítanak meg pleisztocén infúziós lösszel fedett magaslatok. Ez utóbbiakra a mészlepedékes, különösen termékeny csernozjom talajok jellemzők, míg az alacsonyabb térszíneken inkább réti, valamint öntéstalajokkal találkozhatunk (Dövényi 2010).



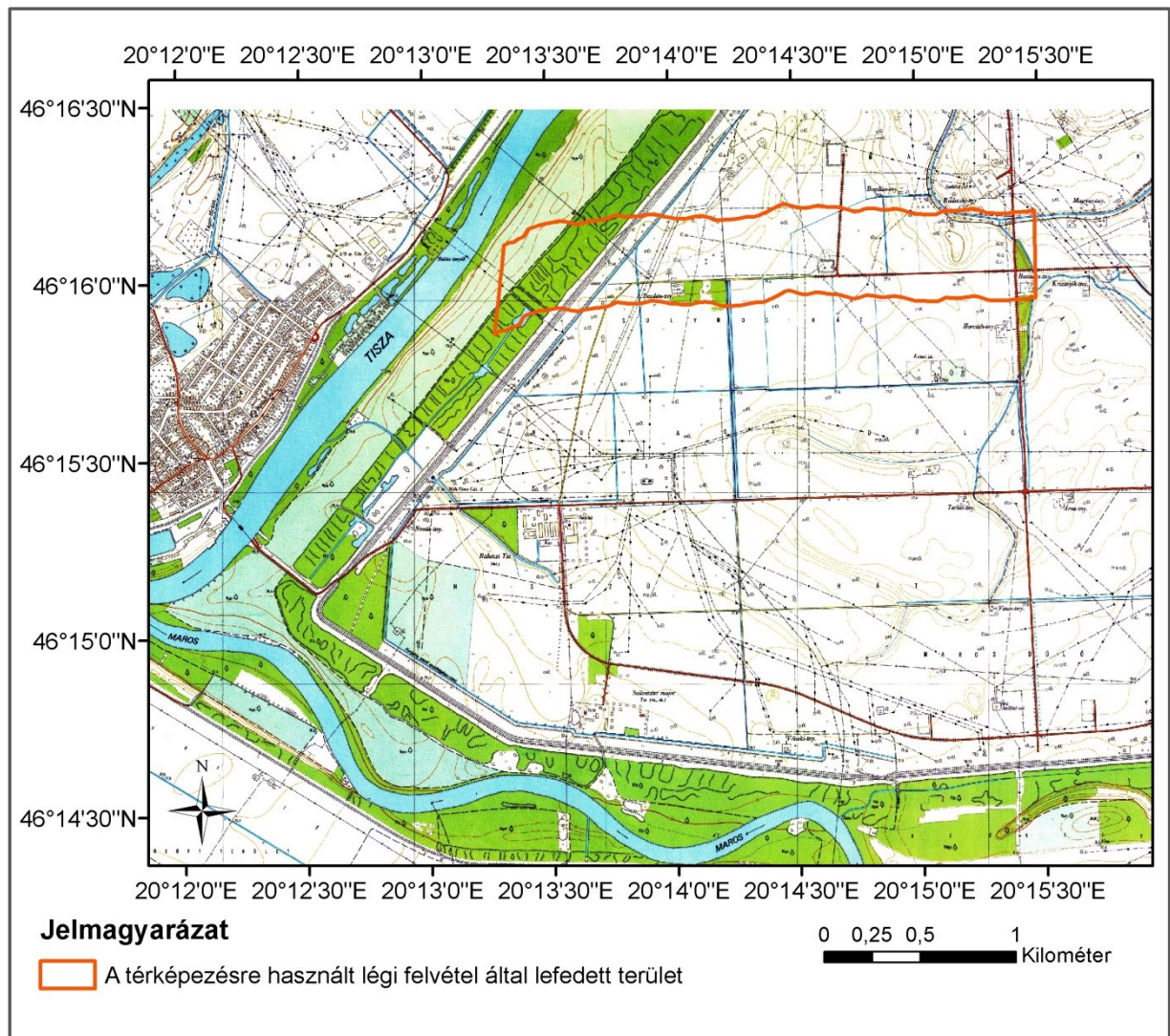
5. ábra: A Tápai-rét és környéke az I. katonai felmérés térképszelvényén (SZTE Földrajzi Térképtár, EOVS vetületi rendszerbe korrigálva)

A kis szemcseméretű, agyagos, ezért a felszíni vizek beszívargását visszatartó talaj – párosulva a magas talajvízszinttel, a folyók közelségével – jelentős belvíz-veszélyeztetettséget idéz elő. Bár a Tápairéten és környékén több szivattyútelep is található, a vízelvezető csatornahálózat karbantartásának elmaradása és a nem körültekintő mezőgazdasági hasznosítás, a kialakult eketalpréteg miatt évről évre időszakos vízelöntésekkel kell számolni.

A belvíz mellett a térség másik komoly környezeti problémája az özönnövények gyors terjedése és az őshonos, természetes vegetáció visszaszorulása. A hullámtéren az amerikai kőris és a gyalogakác kiszorítja a puhafás növénytársulásokat, a hazai folyók mentén megszokott fűz-nyár galériaerdőket (*Saliceto-Populetum*), míg a mentett oldalon a mélyebb, vizenyős területek és a csatornák vannak kitéve az invazív fajok, köztük a nád (*Phragmites australis*) agresszív terjeszkedésének.

Az invazív fajok gyors terjedése összefüggésben van a tradicionális vidéki életforma, és a hozzá kapcsolódó földhasználat változásával. Az egykori rendkívül sűrű tanyavilág alig

egy évtized leforgása alatt néptelenedett el: a termelőszövetkezetek tagosítása 1962-ben kezdődött, a folyamatot az 1970-es árvíz pusztítása tetőzte be (ifj. Lele 1978). Napjainkra már csak néhány tucat család maradt az egykori, több, mint ezer fős külterületi népességből. A táj mai arculatát a nagytáblás szántóföldi művelésen túl a hatvanas években feltárt algyői szénhidrogén-mezőkhöz kapcsolódó ipari műtárgyak, gáz- és olajkutak határozzák meg.



6. ábra: A tájtérképezés mintaterülete, 1:10000-es EOTR térképszelvényen (1984, SZTE Földrajzi Térképtár)

3.2 Adatok

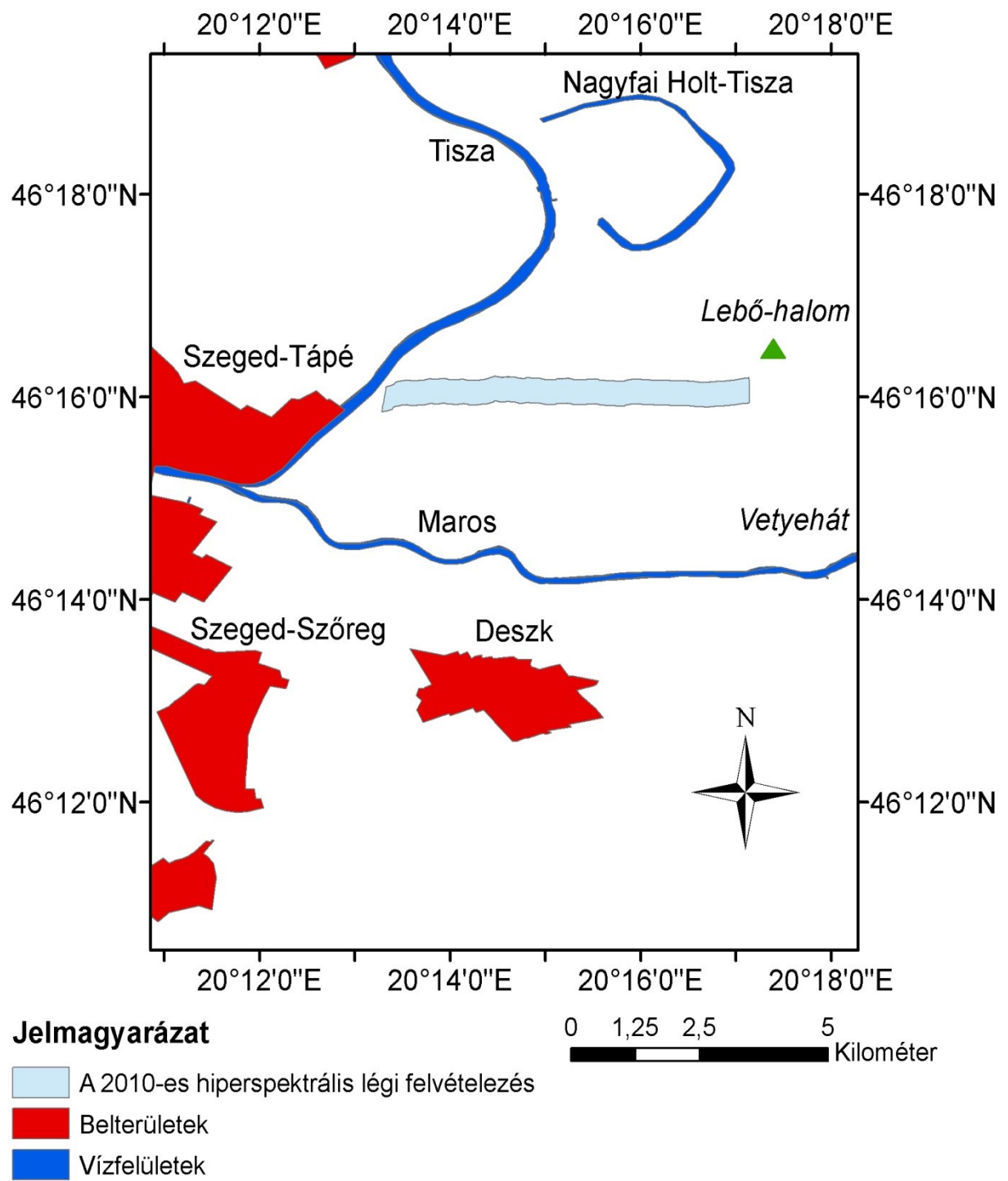
3.2.1 Hiperspektrális adatok

Kutatásaim alapjául a Tápairétről és környékéről 2010. szeptember 22-én készített hiperspektrális képsorozat szolgál (7. ábra, 2. melléklet). A felvételezett pászta hosszúsága 4958, szélessége 640 méter.

A Specim gyártmányú AISA Hawk és Eagle kamerák kombinációjával összeállított AISA Dual hiperspektrális képalkotó rendszer 400 és 2450 nanométeres hullámhossz között rögzíti a felszínről visszaverődő fényt. A 359 spektrális sáv szélessége 4,29 és 6,28 nanométer között változik, a felvételezett adat pedig 14 bites terjedelemben kerül tárolásra.

A képek délelőtt 10:45 és 11:47 között készültek, az ideálisnak tekinthető időszámban, amikor a reggeli, gyors felmelegedés és a kora délutáni hőmérsékleti maximum okozta légköri optikai anomáliák mérsékeltebben jelentkeznek. A felvételezést sem felhőképződés, sem élénkebb szél nem zavarta, a hőmérséklet a repülés ideje alatt 17-20 °C körül alakult. Az Országos Meteorológiai Szolgálat adatbázisa szerint a vizsgálatot megelőző egy hétben jelentős mennyiségű eső hullott a területre. A 2010-es év egészéről elmondható, hogy kivételesen csapadékos volt, az országos átlag elérte a 959 mm-t (OMSZ), ez vezetett a kiterjedt belvízképződéshez is.

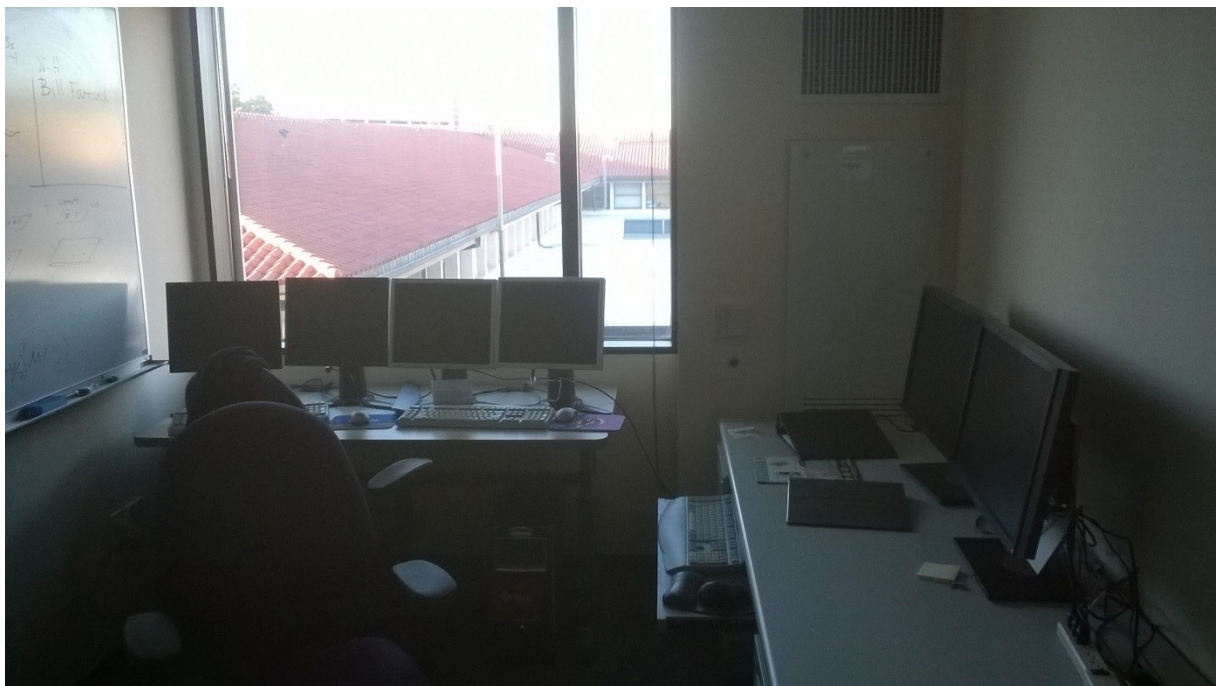
Az adatok ún. push-broom technikával kerültek rögzítésre, melynek lényege, hogy az egymás mellett ülő szenzorok párhuzamosan felvételeznek. Mivel a szenzorok kalibrációja nem tökéletes és az érzékenységükben lehetnek apró eltérések, melyek képi zajként jelentkezhetnek, ezért az adaton Cross-track illumination korrekcióra került sor. A többi, atmoszférikus eredetű zajt az érintett spektrális csatornák eltávolításával szűrtem ki.



7. ábra: A 2010-es hiperspektrális légi felvételezés területe

3.2.2 Adatfeldolgozási környezet

A hiperspektrális adatok feldolgozását két helyszínen végeztem, egyrészt az SZTE Természeti Földrajzi és Geoinformatikai Tanszékének fotogrammetriai laboratóriumában, másrészt a houstoni Rice Egyetem statisztikai tanszékén, a Campus Hungary ösztöndíj-program támogatásával. A szegedi intézetben zajszűrővel és az adatok előfeldolgozásával foglalkoztam, továbbá osztályozásokat és statisztikai kiértékeléseket hajtottam végre az ENVI IDL programban. Az előkészített felvételeket a houstoni képfeldolgozó laboratóriumban vizsgáltam tovább, itt került sor a kereszt-validációs pontthalmazok automatizált kiosztására, a képek normalizációjára és az SOM-hibrid osztályozásra. Az adatkonverzió során nehézséget jelentett, hogy a Rice Egyetem UNIX-alapú rendszerei fordított bájtrendet használnak, ezen kívül a szöveges paraméteres fájlokat is kompatibilis formátumba kellett átalakítani. A műveletekhez a linux operációs rendszerek alatt használatos Bourne Again Shell (BASH) felhasználói felületet alkalmaztam, a szükséges parancsokat és scripteket is ezen a programozási nyelven írtam.



8. ábra: A statisztikai tanszék (Rice) képfeldolgozó laboratóriuma

3.3 Képfeldolgozási módszerek

3.3.1 A hiperspektrális felvételezés alapjai

A távérzékelés a tudományos vizsgálatok alapvető eszköze, amit a műszaki és természettudományok széles köre használ. Alkalmazását leginkább a spektrális felbontás korlátozza: a legtöbb szenzor érzékelési tartománya túl széles a képi információ pontos osztályozásához. A részletesebb radiancia adatok kinyerésére fejlesztették ki a képalkotó spektrométereket, melyeket kezdetben katonai célokra használtak fel, elsősorban a felderítésben (van der Meer et al. 2001)

Az első, polgári célokra használt spektrométert a Geophysical Environmental Research Company-nél fejlesztették ki, a berendezés a 400-2500 nanométerig terjedő spektrális tartományban felvételezett összesen 576 sávban, légi adatgyűjtésre először 1981-ben használták (Chiu & Collins 1978, van der Meer et al. 2001). Az AIS (Airborne Imaging Spectrometer) már csupán 128 adatsatornával rendelkezett, 0,4- 1,2, valamint 1,2 és 2,4 mikrométer között, 1,9 mrad szélességű pillanatnyi látómezeje a szokásos, 4200 méteres repülési magasságból 8 méteres geometriai felbontást tett lehetővé (Lillesand et al. 2004). Az AIS rendszer felváltására 1983-ban a NASA Jet Propulsion laboratóriuma kifejlesztette az AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) rendszert, mely 1987-től szolgáltatja a 400 és 2500 nanométer közötti spektrális információt összesen 224 sávban, átlagosan 10 nanométeres intervallumokban (Green et al. 1998). Napjaink legfontosabb hiperspektrális felvételezési rendszerei között említhetjük még a teljesség igénye nélkül a kanadai CASI, a német DAIS, a finn AISA, a műholdról felvételező Hyperion és a 2010-es évek végén a világűrbe telepítendő EnMAP szenzorait.

A szenzorra érkező jelet különféle sugárosztókkal, szűrővel, prizmával, illetve hiperspektrális képalkotó rendszerek esetén optikai ráccsal bontják szét hullámhossz alapján. A bejövő fény útja egy lencserendszeren keresztül vezet az érzékelőkhöz (van der Meer et al. 2001). A szenzor-tömbben szinte egyidejűleg, ráadásul bármilyen mozgó alkatrész közreműködése nélkül, a lehető legrövidebb idő alatt történik a mérés. Az elektromágneses

sugárzás keltette elektromos töltés félvezetőkön keresztül jut a jelerősítőhöz, majd ezután kerül sor az analóg jel digitalizálására (Platt & Stutz 2008).

A digitális értéket a következő képlettel számíthatjuk át radianciává:

$$L = a + b * DN$$

ahol L a szenzornál mért radiancia, „ a ” a minimális radianciát, azaz az offset értéket jelöli, b a radianciagörbe meredeksége a kiosztott DN digitális értékek tartományában, DN pedig a kalibrált radianciát mutatja, digitális számmal (DN) kifejezve (Markham & Barker 1986 nyomán).

3.3.2 A tanítóterületek kezelése

Távérzékeléssel nyert, raszteres állományú adatoknál megfigyelhető jelenség a területi autokorreláció (spatial autocorrelation), a szomszédos pixelek értékei között fennálló, rendszerint pozitív előjelű korreláció. A szoros kapcsolat ronthatja az osztályozás pontosságát, mivel ebben az esetben a spektrális sávok között is a vártnál erősebb korreláció jelentkezik (Mather 2004). Megoldást a tanítóterületek diverzifikálása, egymástól távolabb eső pontok összeválogatása jelenthet.

Landsat MSS multispektrális műholdfelvételek esetén megfigyelték, hogy véletlenszerűen összeválogatott mintapontok esetén jóval magasabb variancia értékek születtek, mint amikor összefüggő tömbben kerültek az egyes osztályok tanítóterületei kijelölésre (Campbell 1981). A mintavételi területek közötti távolság meghatározásánál érdemes figyelembe venni a területi autokorreláció mértékét is (Labovitz & Matsuoko 1984).

Képosztályozások kiértékelésére leginkább az úgynevezett tévesztési mátrixok (angolul error vagy confusion matrix) alkalmazása terjedt el, melyek táblázati sorokba és oszlopokba rendezve, tételesen mutatják az osztályozási címkék kiosztását a referenciapontokon. Az ellenőrzési adatmintákon értelmezett pontosság (Producer's Accuracy) és az osztályozás megbízhatóságának (User's Accuracy) egymást kiegészítő vizsgálata, ami a rosszabbul teljesítő tanítópontok kiszűrését segítheti, *Story és Congalton (1986)* nevéhez köthető. Az eredménytérkép pontosságát legjobban tükröző statisztikai érték a teljes pontosság (Overall Accuracy), melyet a mátrix főátlójának összege és a mintaszám hányadosaként kapunk meg (Congalton & Green 2009).

A társadalomtudományoktól átvett Cohen-féle Kappa-index képletében az egyezések várható száma is szerepel (Cohen 1960), a módszernek számos módosított változatát használják távérzékelési vizsgálatokban (Congalton & Green 2009).

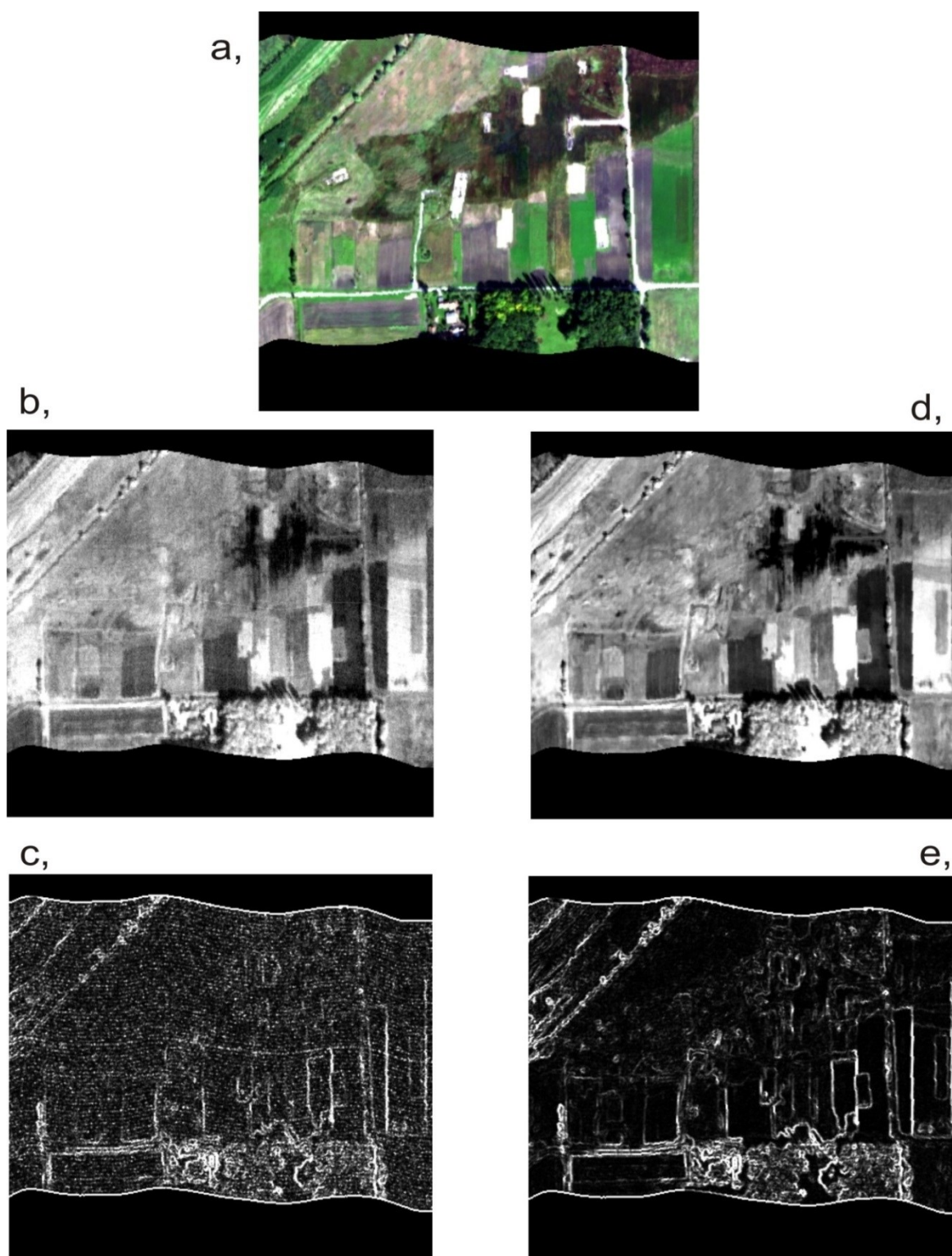
4. EREDMÉNYEK

4.1 Az adatok előfeldolgozásának eredményei

Az osztályozásokhoz és a további adatfeldolgozási lépésekhez elvégeztem a szükséges zajkorrekciós és adatelőkészítő műveleteket, a 359 spektrális csatornából összesen 273-at használtam fel a vizsgálataimban, melyeket a Geary C-féle lokális területi autokorrelációs statisztika alapján válogattam ki, az adathiányos területeket pedig kimaszkoltam. Munkámat részben az SZTE fotogrammetriai laboratóriumában folytattam, ENVI IDL és ERDAS IMAGINE képfeldolgozó programokkal, másrészt a houstoni Rice Egyetem Khoros rendszerét vettem igénybe különféle képi statisztikák kiszámítására.

4.1.1 Zajsűrés

A hibás, ill. zajos sávokat a Geary-féle C mutató (Geary 1954) értékei alapján válogattam ki. Az indexértékekkel az egyes képelemek és szomszédaik közötti, ún. területi autokorrelációt jellemeztem. A 9. ábra a, jelű képén a hiperspektrális légifelvétel egy nyugati kivágata látható valószínűségi kompozícióban, a b, és d, képek ugyanezen terület reflektanciáját ábrázolják a 981 illetve az 1044 nanométeres spektrális sávokban, míg a c, és az e, pontokban az említett két csatorna adataira számított autokorrelációt szemlélteti az ábra, világosabb színnel jelölve a magasabb indexértékeket. A pánkromatikus felvételeiken szinte teljesen megegyező képet mutató adatsávok Geary-féle C autokorrelációs értékei lényegesen eltérnek, a 981 nanométeres spektrális csatornán jelentős zajra utalnak a pásztázás vonalában megjelenő fehér foltok. A feltehetően légköri eredetű zaj ebben az esetben jelentősen befolyásolná a képosztályozások eredményességét, ezért megállapíthatjuk, hogy a spektrális sávok szűrésére a szürkeárnyaltos felvételek vizuális kiértékelésén túl a pixelértékek közötti korreláció vizsgálata is szükséges. A kiválasztott, zajmentes csatornák listája a 3. mellékletben található.



9. ábra: A hiperspektrális felvétel valószínűségi kivágata (a), a 981 nanométeres sáv szürkeárnyaltos képe (b), a 981 nm-es sáv Geary-féle C területi autokorrelációs térképe (c), az 1044 nm-es sáv szürkeárnyaltos képe (d) és az 1044 nm-es sáv Geary-féle C területi autokorrelációs térképe (e). A c, és az e, képeken világos szín jelzi a magasabb értékeket.



















4.1.2 Adatnormalizáció

A neurális hálózat gyorsabb tanulása érdekében érdemes a bemenő adatokat azonos értéktartományban kezelni. Erre a célra egyfajta lineáris kontrasztosítást (linear contrast stretch) alkalmaztam, hasonlóan *Skidmore et al. (1997)*, *Richards (1986)* és *Knox et al. (2011)* adatnormalizációs eljárásához. Meghatároztam a felvételen rögzített legmagasabb és legalacsonyabb reflektancia értéket, valamint a kép szegélyén található, hiányos (no data) képelemek számkódját és ezek alapján transzformáltam át a reflektancia adatokat.

4.2 Képi statisztikai vizsgálatok és adatredukciós eljárások

4.2.1 A vizsgált osztályok

A terepi felmérések során összegyűjtött és a légi felvételeken is azonosított osztályok átlagos spektrumgörbéit mutatja be a 1. melléklet. Képi statisztikáimat és az osztályozásokat ugyanazokon a tanítóterületeken végeztem, ezek jellemzően 40-60 mintát tartalmaznak. A talajfelszínek, a természetes és a mezőgazdasági növényfajok mellett három mesterséges anyagot is azonosítottam a területen (aszfalt, beton, tetőcserép). A terepi információk összegyűjtését a képosztályozások fejezeteiben részletezem, az osztályok pontos színbeosztása és betűkódjai a 10. ábrán láthatók. A következőkben az adatredukció (feature reduction) elméleti hátterét mutatom be, majd az erre épülő statisztikai eljárásokat a kijelölt osztályokon alkalmazom.

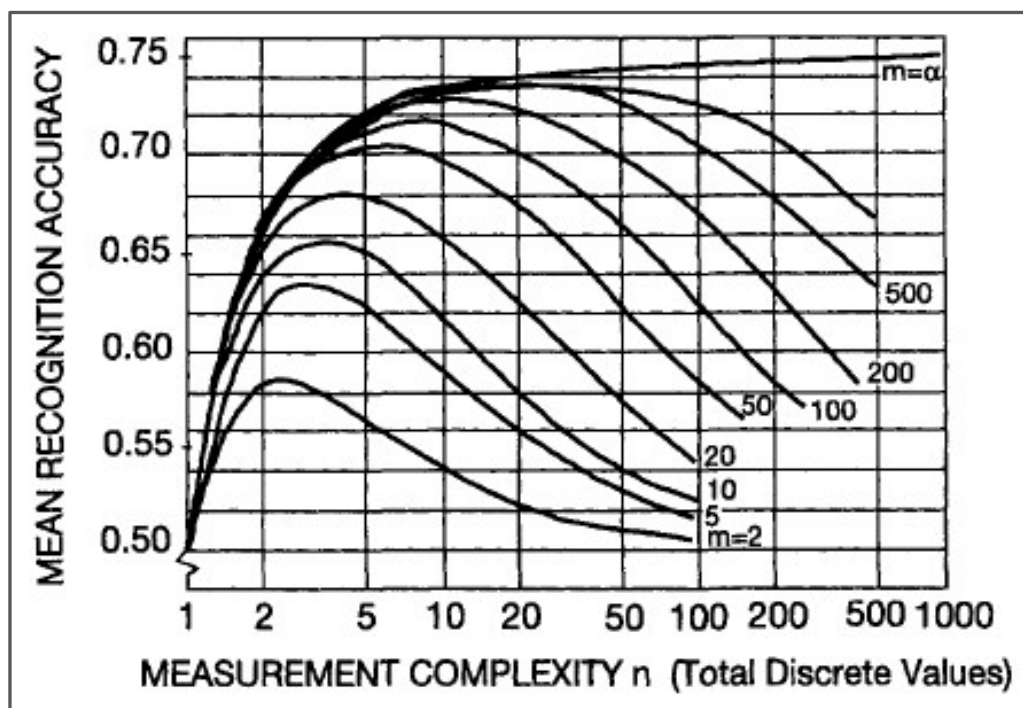
Osztályok		
 Amerikai kőris (A)	 Gyalogakác (G)	 Vízben álló növényzet (U)
 Aszfalt (B)	 Gyepfoltok (H)	 Nyílt vízfelület (V)
 Beton (C)	 Nemesnyaras (J)	 Lucerna (W)
 Bolygatatlan talajfelszín (D)	 Fehér fűz (K)	 Kukorica (X)
 Bolygatott talajfelszín (E)	 Szürke nyár (M)	 Fokhagyma (Y)
 Piros tetőcserép (F)	 Tölgy (N)	 Cukorrépa (Z)

10. ábra: A vizsgált osztályok kutatásomban használt szín- és betűkódjai

4.2.2 A Hughes-hatás

A Hughes-hatásnak vagy Hughes-jelenségnek (Hughes Effect, Hughes Phenomenon) nevezett, eredetileg elméletben levezetett statisztikai törvényszerűség az adat dimenzionalitása és az osztályozási pontosság közötti összefüggést írja le, ami konstans, véges számú tanítópont esetén az adatkomplexitás növekedésével emelkedő, majd egy pont után enyhén csökkenő mintázatot mutat (Hughes 1968).

A Hughes-hatás jellegzetes pontossági görbéje, két függvény, két, első ránézésre egymásnak látszólag ellentmondó statisztikai jelenség eredőjeként áll elő (Landgrebe 2003). Egyrészt, ahogy szélesedik az adatok értelmezési tartománya, esetünkben a spektrális sávok száma, úgy javul az egyes minták elkülöníthetősége, csökken az osztályok közti átfedés és javul a klasszifikációs pontosság. Másrészt viszont, a dimenziók számának emelkedésével egyre nagyobb adattérre kell a statisztikai számításokat kiterjeszteni, ami, ha a tanítópontok körét nem bővítjük, végső soron az osztályozási pontosság csökkenéséhez vezet.



11. ábra: Az adatkomplexitás és a pontosság összefüggése

(Landgrebe 2003, Hughes 1968)

A Hughes-jelenség egyik fontos tényezője a minták számossága. Egy hipotetikus, végtelen számú tanítóelemet tartalmazó osztályozás esetében ugyanis a pontosság a dimenzionalitás bővülésével monoton növekszik. A pontosság csökkenését csak véges számú mintáknál figyelhetjük meg, ráadásul a görbe maximumának helye (és értéke) függ a mennyiségüktől is: több tanítópont esetén magasabb adatkomplexitásnál találjuk a függvény optimumát és a pontossági arány is nagyobb (11. ábra). A „dimenzionalitás átka”-ként is emlegetett anomáliát tehát úgy is értelmezhetjük, hogy a vizsgálatok összetettségét követnie kell a tanítók számának.

A tanítóterület mérete azonban a legtöbb távérzékelési vizsgálat esetén korlátozott (Serpico et al. 2007), ezért az osztályozási pontosság optimalizálására más módszert kell alkalmazni, a bemeneti, spektrális sávok adattartalmát, adat-komplexitását csökkenteni kell (Shaw & Manolakis 2002). A „Feature Reduction” (adatredukció) gyűjtőnévvel illetett módszerek közé tartozik a sávok szelekciója, valamint a különféle, lineáris és nemlineáris transzformációk is (Serpico et al. 2007).

4.2.3 Az adatredukció módjai

A Hughes-hatás következményeként a spektrális adatmennyiség növekedése a pontosság romlását vonja maga után (Hughes 1968), a nagyobb felbontású felvételek számítógépes feldolgozása pedig jelentős memória kapacitást igényel, hosszabb feldolgozási idővel. Ha nincs lehetőség elégséges számú tanítópont kijelölésére, akkor az adat dimenzionalitását, azaz a vizsgálatba vont spektrális csatornák körét kell szűkíteni. Az angolul Feature reduction-nek nevezett műveletre két lehetőség kínálkozik: a képi adatot transzformálhatjuk egy alacsonyabb dimenzionalitású adattérbe (Feature Extraction), vagy egyszerűen kiszűrhetjük a kevésbé fontos adatsávokat (Feature Selection). Előbbire jó példa a főkomponens analízis és az MNF transzformáció, több száz sávban felvételezett hiperspektrális képeknél azonban célravezetőbb és kevesebb számítási kapacitást igénylő megoldás lehet a zajos, redundáns, illetve egymással erős korrelációt mutató adatsávok mellőzése. Mindkét megközelítésnél fontos szerepe van az adatok hasznos információtartalmát értékelő szeparabilitási indexeknek.

4.2.4 Főkomponens analízis

Az adattömörítés és –transzformáció egyik lehetséges módja főkomponens sávok készítése (Principal Component Analysis). Az új adattér meghatározásához elsőként a felvételek kovariancia vagy korreláció mátrixát kell kiszámítani (utóbbit akkor célszerű alkalmazni, ha a képelemek a különböző spektrális csatornákon lényegesen eltérő értékeket vesznek fel). A mátrix sajátértékeinek segítségével lehet az ideális főkomponens (Principal Component) sávokat kialakítani. A spektrális diszkrimináció szempontjából az alsóbbrendű sávok is értékes információval szolgálhatnak, viszont bizonyos esetekben, főleg hiperspektrális adatoknál jelentős, varianciájában az érdemi képi tartalmat felülmúló zajjal kell számolni (Tso & Mather 2009).

A légi felvételen főkomponens transzformációt végeztem kovariancia mátrix alapján, a normalizáció előtti és a már normalizált adatokon is. A hiperspektrális képek jellegzetessége, hogy az adattérben a legnagyobb variáciát gyakran az intenzitásbeli különbségek idézik elő, azaz az első főkomponens csupán a radiancia intenzitását mutatja, a spektrumgörbék eltérő futásából adódó különbségeket a további sávok jelenítik meg (Comelli et al. 2008). A transzformált adatom első sávjában ennek megfelelően a megvilágítás erőssége látszódik, ami a délelőtti felvételezés miatt délkeleti irányból éri a területet (12. ábra). A sajátérték ráták alapján teljes képi információtartalom több mint 99 %-át ez a sáv tartalmazza (1. táblázat). A probléma megoldására két megoldás kínálkozik: vagy figyelmen kívül hagyjuk az első főkomponenst, vagy pedig a transzformációt eleve a megvilágítási anomáliáktól megtisztított, normalizált adaton hajtjuk végre. Ez utóbbi esetben már az első sáv is tartalmaz érdemi spektrális információt és a sajátérték is arányosabban oszlik el a főkomponensek között (1. táblázat). További vizsgálataimban a nem normalizált adat főkomponens sávjait használtam.

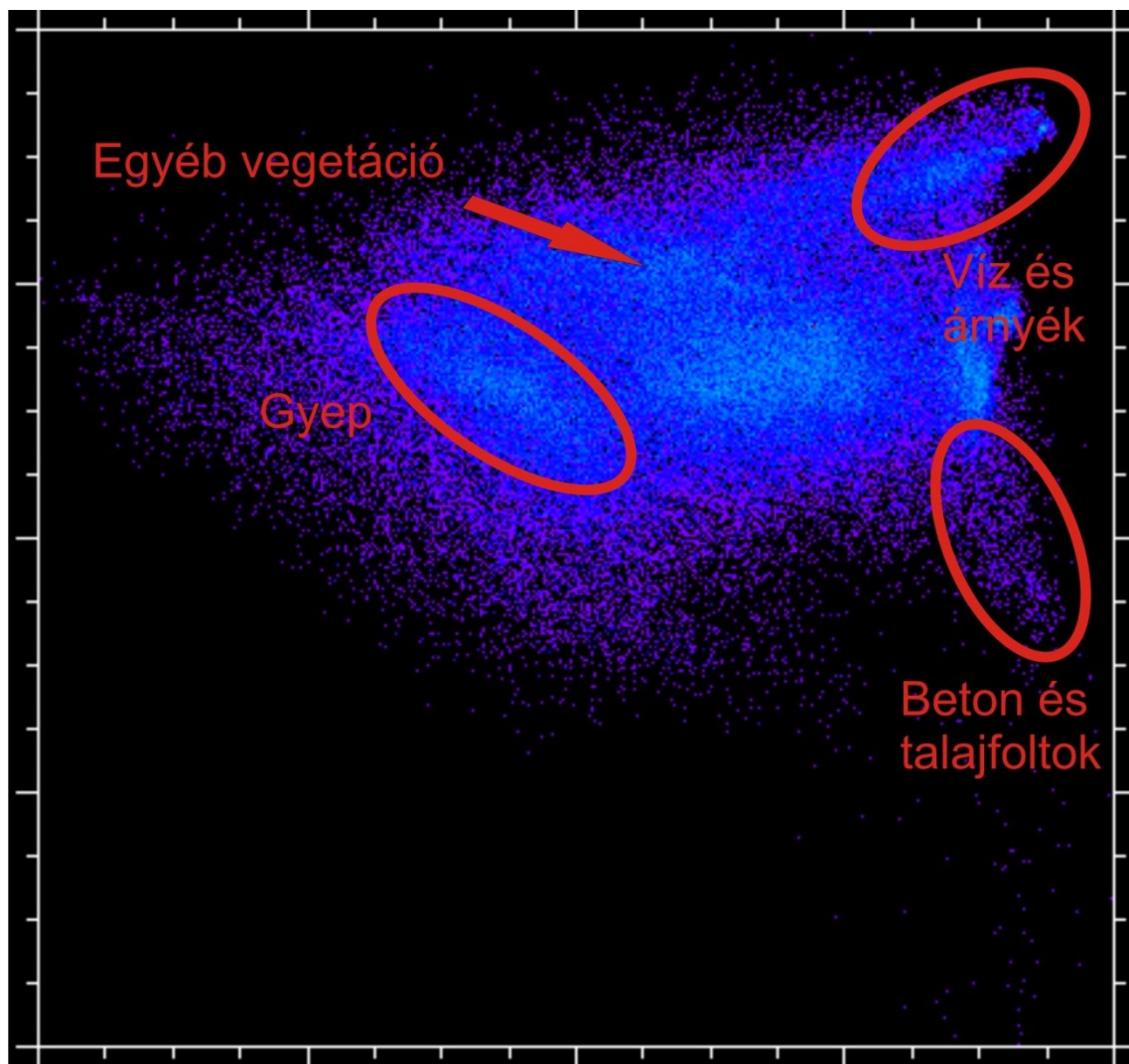
PC sáv	PC1	PC2	PC3	PC4	PC5	PC6
Sajátérték ráta (nem normalizált adaton)	99,25%	0,65%	0,08%	0,01%	0,004%	0,003%
Sajátérték ráta (normalizált adaton)	93,72%	4,74%	1,08%	0,2%	0,07%	0,07%

1. táblázat: A főkomponens sávok, és sajátérték rátájuk



12. ábra: Az első főkomponens sáv szürkeárnyaltos képe

Mintaterületem képi adatait az albedó hatástól mentes első két főkomponens sáv terébe vetítettem ki, az úgynevezett scatter plot diagram a felosztástípusok spektrális hasonlóságait hivatott topológiailag egzakt módon bemutatni (13. ábra). Az azonosított három fő kategória (növényzet, talaj, víz) hármasszoros elhelyezkedése megfelel a spektrális keveredés irodalmában leírtaknak (Mucsi & Henits 2011). A két sáv értékeire kivetített adatfelhő az ugyanezen az adatokon elvégzett szeparabilitási vizsgálat (4.2.5 fejezet) eredményeit tükrözi. A vízfoltok, a talajfelszínek és a gyepel borított területek viszonylag jól elkülönülnek a többi felületborítási osztálytól, keveredést, illetve kevésbé éles átmenetet a növényzet esetén tapasztalhatunk. A sűrűségi színezésen világosabban megjelenő képelemek jeleznek magasabb koncentrációt, ezeket pedig elsősorban a háromszög alakzat vegetáció által dominált, központi részein találjuk (13. ábra).



13. ábra: Képelemek a két főkomponenssáv spektrális terében

4.2.5 Spektrális szeparabilitási számítások

Az osztályok közötti átfedések mérésére használhatók az úgynevezett spektrális szeparabilitási számítások, melyekkel a reflektancia értékek eloszlási függvényeit hasonlítjuk össze. Bár a legtöbb esetben megbízhatóan jellemzik a tanító adatbázisok mintáinak spektrális összefüggéseit, a szeparabilitási vizsgálatok ugyanakkor pontatlan eredményt adhatnak nem normál eloszlású adatsorok esetén, mivel ebben az esetben az egyes osztályokhoz tartozó kovariancia mátrixok kiszámítása önmagában nem elégséges az osztályok közötti kapcsolatok, az esetleges átfedések számszerűsítéséhez (Tso & Mather 2009).

A szeperabilitási értékek azt is kifejezik, hogy a spektrális értékek változékonysága osztályon belül vagy osztályok között nagyobb (Clark et al. 2005), a nagyon alacsony szeperabilitású osztályokat a kutatásokban gyakran összevonják és általánosságban is elmondható, hogy a módszer jól alkalmazható a tanítóterületek kiválasztására és ellenőrzésére (Thomas et al. 2002, Ustin et al. 2009, Metternicht & Zinck 1998).

A Bhattacharyya-féle távolságot az alábbi formulával írhatjuk fel:

$$B = \frac{1}{8} (\mathbf{m}_i - \mathbf{m}_j)^t \left\{ \frac{\Sigma_i + \Sigma_j}{2} \right\}^{-1} (\mathbf{m}_i - \mathbf{m}_j) + \frac{1}{2} \ln \left\{ \frac{|\Sigma_i + \Sigma_j|/2}{|\Sigma_i|^{1/2} |\Sigma_j|^{1/2}} \right\}$$

ahol \mathbf{m} jelzi az osztályokhoz (i es j) tartozó minták átlagát, Σ a kovariancia mátrixukat, T pedig a transzponálásra utal. A Bhattacharyya-féle számítás módosításából áll elő a Jeffries-Matusita index képlete is, melyet a következőképpen áll elő:

$$J = 2 (1 - e^{-B})$$

A képi statisztikában elterjedt divergencia képletének módosításából származik a Transformed Divergence (TD) szeperabilitási függvény, melynek értékei az osztályátlagok közötti különbségekkel exponenciálisan növekednek, alkalmazása a képfeldolgozásban mindenekelőtt szerényebb számítási igényének köszönhetően terjedt el (Richards & Jia 2006). *Jensen (1986)* megállapítása szerint elkülöníthetőség szempontjából az 1,9-es TD index jelenti a választóvonalat, az ennél alacsonyabb értékek jelentős átfedéseket jeleznek.

A tanítóterületek spektrális értékei között átfedések lehetnek, ami megnehezítheti a későbbi osztályozásokat, ezért egy áttekintő jellegű szeperabilitási táblázatot készítettem az egyes osztályok mintapontjainak elkülöníthetőségéről. Mivel a szeperabilitási vizsgálatokra is érvényes az n+1 megszorítás, miszerint a pontok számának meg kell haladnia a spektrális csatornákét, ezért az elemzést a két, illetve öt legnagyobb varianciájú, albedó hatástól mentes főkomponens sávon végeztem el.

	A	B	C	D	E	F	G	H	J	K	M	N	U	V	W	X	Y	Z	TD
A	0	2	2	2	2	2	2	2	1,9	1,19	2	2	2	2	2	2	2	2	A
B	2	0	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	B
C	2	2	0	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	C
D	2	2	2	0	1,93	2	2	2	2	2	2	2	2	2	2	2	2	2	D
E	2	2	2	1,6	0	2	2	2	2	2	2	2	2	2	2	2	2	2	E
F	1,99	1,85	1,83	2	2	0	2	2	2	2	2	2	2	2	2	2	2	2	F
G	1,78	2	2	2	2	2	0	2	1,92	1,99	1,96	1,99	2	2	2	2	2	2	G
H	1,81	2	2	2	2	2	2	0	2	1,99	2	2	2	2	2	2	2	2	H
J	1,64	2	2	2	2	2	1,8	2	0	1,54	1,04	0,8	2	2	2	2	2	2	J
K	1	2	2	2	2	2	1,89	1,87	1,45	0	1,89	1,94	2	2	2	2	2	2	K
M	1,85	2	2	2	2	2	1,85	2	0,87	1,56	0	1,29	2	2	2	2	2	2	M
N	1,79	2	2	2	2	2	1,92	2	0,7	1,82	1,05	0	2	2	2	2	2	2	N
U	1,9	2	2	2	2	2	1,98	2	1,7	1,98	1,93	1,92	0	2	2	2	2	2	U
V	2	2	2	2	2	2	2	2	2	2	2	2	2	0	2	2	2	2	V
W	2	2	2	2	2	2	1,99	2	2	2	1,96	2	2	2	0	2	2	2	W
X	1,98	2	2	2	2	2	2	2	2	1,99	2	2	2	2	2	0	1,96	1,68	X
Y	1,95	2	2	2	2	2	2	2	2	1,98	2	2	2	2	2	1,89	0	1,11	Y
Z	1,97	2	2	2	2	2	2	2	2	1,99	2	2	2	2	2	1,56	1,02	0	Z
JM	A	B	C	D	E	F	G	H	J	K	M	N	U	V	W	X	Y	Z	

2. táblázat: Az osztályok közötti Jeffries-Matusita (JM) és Transformed Divergence (TD) szeparabilitási értékek az első öt főkomponens sáv alapján

Öt sáv képi adatai alapján számított indexértékek találhatók a 2. táblázat soraiban és oszlopaiban, a mátrix átlója mentén 0 jelzi az osztályok teljes átfedését önmagukkal, ennek egyik oldalán, a jobb felső sarokban a Transformed Divergence (TD), a bal alsó sarokban a Jeffries-Matusita (JM) számai láthatók, a legmagasabb lehetséges érték mindkét esetben 2.

A Transformed Divergence módszer kevésbé érzékenyen méri az elkülöníthetőség különbségeit (Jensen 1986), a $18 \times 18 / 2 = 162$ lehetséges kombinációból csupán 17 esetben talált átfedést, a többi esetben az eredmény 2-es értéket adott. Az átfedések a bolygatott (E) és a bolygatatlan (D) talajok közti hasonlóság mellett a legtöbb természetes növényfaj, valamint három mezőgazdasági növény (cukorrépa – Z, fokhagyma – Y, kukorica - X) osztályait érintik. Lényegében ugyanezek az osztályok jelennek meg a Jeffries-Matusita index celláiban 1,7 alatti értékkel (2. táblázat), általánosságban pedig elmondható, hogy a Transformed Divergence módszer alapján jelzett esetekben itt még alacsonyabbak a mutatószámok, különösen nehéz elkülöníthetőségével pedig öt vegetációs osztály tűnik ki, az amerikai kőris (A), a nemesnyár (J), a fehér fűz (K), a szürke nyár (M) és a tölgy (N).

Számításaimat két főkomponens sávon is elvégeztem, melyek ezáltal érzékenyebbek a kisebb mértékű átfedésekre, az öt sávon végzett vizsgálatokban megemlítetteken túl a mesterséges felszínek és bizonyos, spektrálisan jobban elkülönülő növényfajok (pl. gyepek) esetén is adódtak alacsonyabb értékek (3. táblázat). Figyelembe kell azonban venni, hogy a szűkebb adattartományon végrehajtott elemzés kevésbé megbízható, mivel a megkülönböztetés szempontjából fontos képi információk maradhattak ki a számításokból.

A szeparabilitási eredmények alapján az öt főkomponens sávra számított Jeffries-Matusita index kb. 1,7-1,8-as értékeinél húzható meg az a határ, amely alatt szignifikáns, több módszerrel is azonosítható spektrális átfedéssel számolhatunk.

	A	B	C	D	E	F	G	H	J	K	M	N	U	V	W	X	Y	Z	TD
A	0	2	2	2	2	2	0,88	1,9	0,32	0,36	0,88	0,71	1,77	2	2	2	1,97	1,99	A
B	1,84	0	1,95	2	1,98	1,99	2	2	2	2	2	2	2	2	2	2	2	2	B
C	2	1,91	0	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	C
D	1,98	1,62	2	0	0,28	2	2	2	2	2	2	2	2	2	2	2	2	2	D
E	1,98	1,45	2	0,26	0	2	2	2	2	2	2	2	2	2	2	2	2	2	E
F	1,94	1,57	1,34	1,89	1,85	0	2	2	2	2	2	2	2	2	2	2	2	2	F
G	0,73	2	2	2	2	2	0	0,96	0,93	0,81	1,12	1,1	2	2	1,98	2	2	2	G
H	1,35	2	2	2	2	2	0,78	0	1,96	1,94	1,98	1,97	2	2	2	2	2	2	H
J	0,29	2	2	2	2	2	0,82	1,64	0	0,21	0,36	0,22	1,79	2	1,99	2	1,74	1,9	J
K	0,35	1,98	2	2	2	1,99	0,6	1,36	0,19	0	0,2	0,36	1,98	2	1,99	2	1,97	2	K
M	0,81	2	2	2	2	2	1,04	1,77	0,35	0,18	0	0,25	1,99	2	1,96	2	1,9	1,97	M
N	0,54	1,98	2	2	2	1,99	1,04	1,83	0,2	0,29	0,24	0	1,92	2	1,97	2	1,86	1,95	N
U	1,01	2	2	2	2	2	1,88	1,99	1,13	1,34	1,5	1,13	0	2	2	2	1,95	1,95	U
V	1,98	1,98	2	2	2	1,96	2	2	2	2	2	1,99	2	0	2	2	2	2	V
W	1,88	2	2	2	2	2	1,71	2	1,89	1,59	1,52	1,62	2	2	0	2	2	2	W
X	1,71	2	2	2	2	2	1,98	2	1,45	1,46	1,28	1,53	1,99	2	2	0	1,08	0,58	X
Y	1,24	2	2	2	2	2	1,83	2	0,87	1,12	1,03	0,93	1,72	2	2	0,78	0	0,27	Y
Z	1,42	2	2	2	2	2	1,92	2	1,06	1,26	1,11	1,1	1,81	2	2	0,53	0,26	0	Z
JM	A	B	C	D	E	F	G	H	J	K	M	N	U	V	W	X	Y	Z	

3. táblázat: Az osztályok közötti Jeffries-Matusita (JM) és Transformed Divergence (TD) szeparabilitási értékek az első két főkomponens sáv alapján

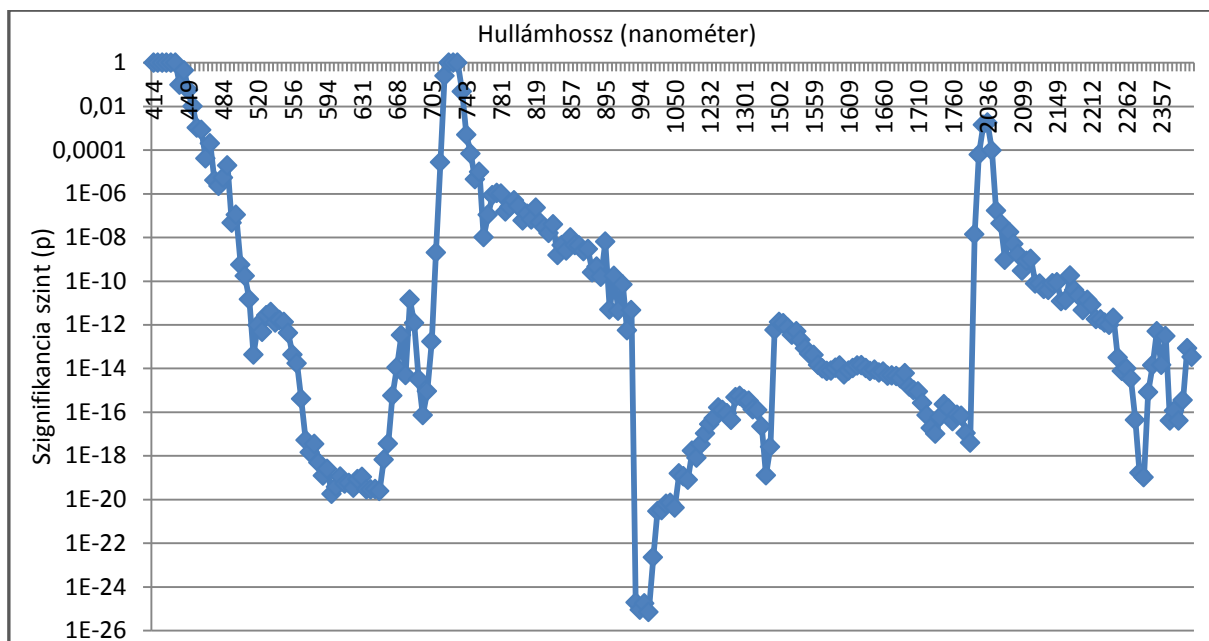
4.2.6 Spektrális diszkrimináció sávonként

Nagy spektrális felbontású képi adatoknál lehetőségünk van a vizsgált osztályok elkülönítésére legalkalmasabb csatornák kiválogatására, melyek rendszerint az elnyelési sávokkal és a reflektancia görbék egyéb jellegzetes, a többi osztályétól eltérő futású

szakaszaival esnek egybe. Matematikai-statisztikai eljárásokkal, függvényvizsgálatokkal, valamint regressziós számításokkal lehet számszerűsíteni az elkülöníthetőség mértékét az egyes spektrális sávokon, az így kapott indexértékeket azonban fenntartásokkal érdemes kezelni, mivel nem feltétlenül tükrözik a kiválasztott legmegfelelőbb sávok tényleges osztályozási eredményeit (Visser et al. 2013), egyes machine learning eljárások példája pedig azt mutatja, hogy magasabb pontosság érhető el, ha valamennyi, a sávselekciónál kimaradt, zajmentes adatsávot felhasználjuk az elemzésekben (Howell et al. 1994, Moon et al. 1995, Merényi et al. 2014).

Növényfajok esetén is jelentős szerepük lehet az abszorpciós sávoknak az adatminták elkülönítésében, gyakran az egyedi biokémiai komponensekhez köthető spektrális anomáliák adják a legfontosabb támpontokat a diszkriminációhoz (Schmidt & Skidmore 2003). Szeparabilitási és egyéb statisztikai vizsgálatok (T-próba, Mann-Whitney-féle U-próba) alapján jellemzően a látható és a közeli infravörös tartományban található a megfelelő sávok (Visser et al. 2013, Adam & Mutanga 2009).

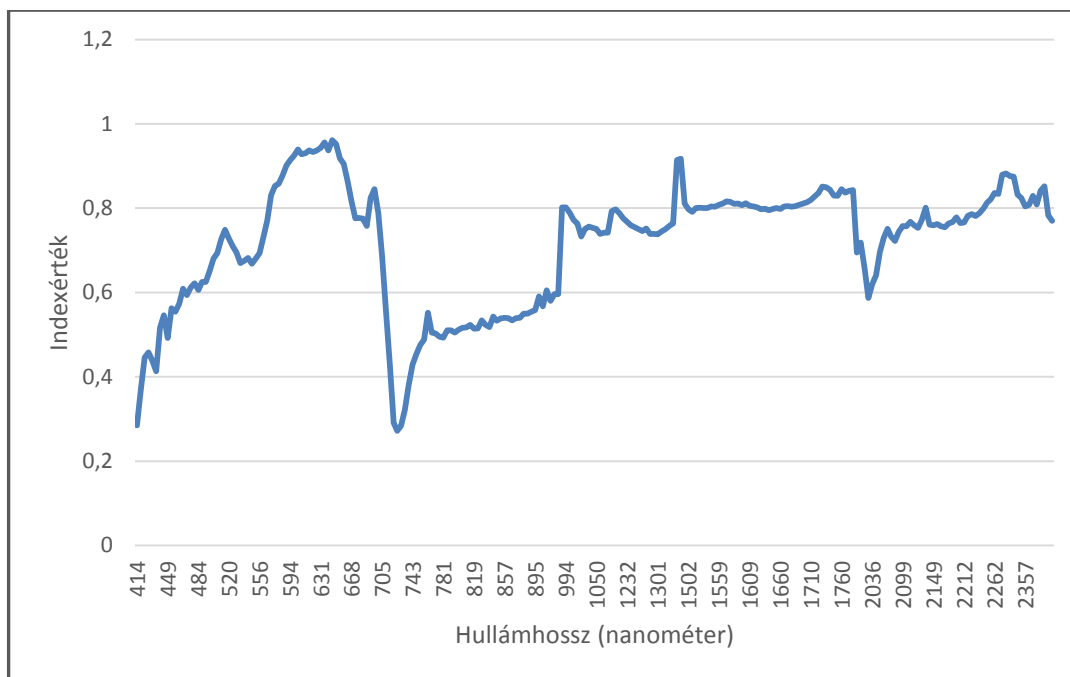
Kísérleti elemzésemben a gyalogakác és a szürke nyár 60-60 darab spektrális mintájának statisztikai elkülöníthetőségét számszerűsítettem adatsávonként, a mérésre egy statisztikai hipotézisvizsgálatot (T-próba) és a spektrális szeparabilitási távolságot használtam. Kutatásaim a 359 hiperspektrális csatorna közül csak arra a 240-re terjedtek ki, melyek zajtartalma még elfogadható szinten volt a térképezésekhez. A kétmintás T-próba esetén a szignifikancia (p) alacsonyabb értékei (14. ábra), míg a Jeffries-Matusita szeparabilitási index nulla és kettő közé eső mutatószámai közül a magasabbak jelzik azokat a spektrális sávokat, ahol a két növényfaj reflektanciája jobban elkülönül (15. ábra). A két módszer eredményei, bár nagyon hasonló futású görbéket adnak, melyek lényegében egymás tükörképei, mégsem egyeznek meg teljesen, a sávok egymáshoz képest felállított, relatív rangsorában kisebb eltérések állapíthatók meg a két adatsor között (5. melléklet).



14. ábra: Kétmintás T-próba szignifikancia (p) szintjei a szürkenyár és a gyalogakác spektrális mintáin, nem egyenlő varianciát feltételezve, a sávok száma alapján korrigálva

A kapott szignifikancia és indexértékek összecsengenek a más szerzők által már leírtakkal (Visser et al. 2013, Adam & Mutanga 2009), amennyiben az 5-600, illetve a 900-1000 nanométer közötti spektrális csatornákat jellemzik magasabb elkülöníthetőségi értékekkel. A látható fény kék tartományának alsó felén (400-450 nm) kívül a vegetációs vörös él régiójában (700-750 nm) tapasztalható jelentősebb átfedés a két osztály között, kedvezőtlenebb indexszámokat ezeken kívül a 2000 nm körüli, légköri zajjal terhelt sávokban találunk.

A vártnál jobb az elkülöníthetőség a magasabb, rendszerint zajosabb, 1000 nanométer fölötti hullámhosszokon, ezért az itt felvételezett spektrális sávok felhasználásától az osztályozási pontosság javulását várhatjuk. A vizsgált felszíntípusok felismerésére legalkalmasabb sávkombinációkat a továbbiakban automatizált keresőalgoritmusokkal találhatjuk meg, eredményüket pedig vegetációs indexek, spektrális indikátorok fejlesztésén túl a légi felvételezés megtervezésénél is alkalmazhatjuk.



15. ábra: A szürkenyár és a gyalogakác Jeffries-Matusita módszerrel számított szeparabilitási értékei, spektrális sávonként

4.2.7 A spektrális homogenitás vizsgálata

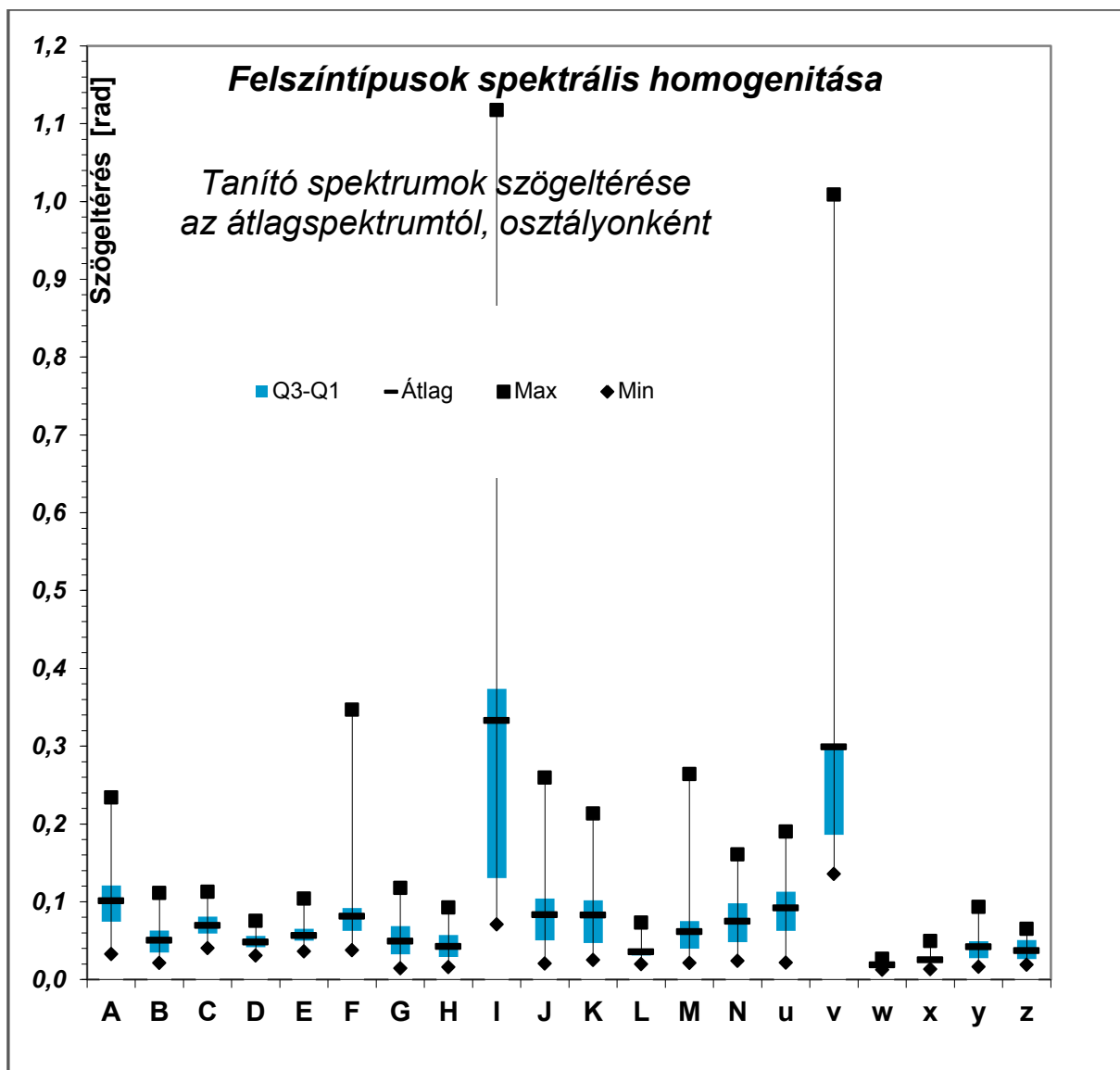
Az osztályok homogenitásának jellemzésére minden ötvenedik zajmentes spektrális sávra kiszámítottam a mintapontok képértékeinek szórását (a 273 elérhető csatornából, 3. melléklet). Nem meglepő, hogy a legkisebb változékonyságot a víz (v) reflektanciája mutatja (legnagyobb szórása a 100. sávban van, 0,011), a sávokhoz tartozó legmagasabb értékeket pedig általában az összetettebb mesterséges felszíneknél, mindenekelőtt a tetőcserepek sorában (F) találhatjuk (4. táblázat). A magasabb szórású, nem mezőgazdasági növényfajok közül valamennyi elemzett adatsávban kiemelkednek az amerikai köris mutatói, különösen szembetűnő a látható tartományban (50. sáv) mért kiugró értéke (0,023), amit az Észak-Amerikából betelepített fák korai, szeptember végén már jelentkező sárgulásával, a lombkorona szélesebb színskálájával magyarázhatunk (a légi felvétel szeptember 22-én készült). A statisztikai adatsorokon további fontos spektrális jellegzetességeket fedezhetünk fel, mint a talajfelszínek alacsony spektrális változékonysága vagy a vegetáció szórásértékeinek ugrásszerű növekedése a látható tartomány fölött, ennek ellenére a kapott számok a teljes felvételezési spektrumnak csak a töredékét reprezentálják. Számításba kell

venni azt is, hogy az eredményeket nagyban befolyásolják az intenzitásbeli eltérések, azaz a megvilágítás erőssége, ezért olyan új eljárást kerestem, ami megoldást kínál az említett módszertani dilemmákra.

Osztály	Band 50	Band 100	Band 150	Band 200	Band 250
Amerikai kőris (A)	0,0229	0,1126	0,1033	0,0610	0,0323
Aszfalt (B)	0,0233	0,0469	0,0510	0,0501	0,0461
Beton (C)	0,0840	0,0974	0,0722	0,0638	0,0585
Bolygatatlan talaj (D)	0,0053	0,0124	0,0121	0,0101	0,0074
Bolygatott talaj (E)	0,0051	0,0167	0,0175	0,0158	0,0081
Tetőcserép (F)	0,0374	0,1224	0,1776	0,1750	0,1509
Gyalogakác (G)	0,0049	0,0770	0,0595	0,0279	0,0115
Gyep (H)	0,0032	0,0764	0,0641	0,0154	0,0084
Nemesnyár (J)	0,0079	0,0683	0,0710	0,0336	0,0134
Fehér fűz (K)	0,0123	0,0877	0,0894	0,0441	0,0213
Szürke nyár (M)	0,0101	0,0843	0,0810	0,0380	0,0156
Tölgy (N)	0,0074	0,0968	0,0895	0,0379	0,0128
Vízben álló növény (u)	0,0035	0,0239	0,0316	0,0217	0,0111
Nyílt víz (v)	0,0020	0,0111	0,0090	0,0065	0,0046
Lucerna (w)	0,0034	0,0384	0,0276	0,0183	0,0090
Kukorica (x)	0,0021	0,0240	0,0141	0,0051	0,0048
Fokhagyma (y)	0,0032	0,0371	0,0192	0,0061	0,0059
Cukorrépa (z)	0,0029	0,0290	0,0157	0,0062	0,0053

4. táblázat: Szórásértékek osztályonként az 50., 100., 150., 200. és 250. sávokban

A tanítótérületek spektrális homogenitásának jellemzésére olyan eljárást kívántam alkalmazni, ami a több száz hiperspektrális sávból álló képi adatok statisztikai leírására is használható. A szeparabilitási számítások és a spektrális szögeltérés-alapú indexek ugyan jól mutatják az osztályok átlagspektrumai közti hasonlóságokat (Tobak et al. 2012a), a tanítótérületen belüli homogenitás mérésére azonban alkalmatlanok. Box plot diagramok készítéséhez egy dimenzióra kell redukálni a nagyfelbontású adatot, amihez a főkomponens transzformáció kínálkozna megfelelő eljárásnak, a hasznos információ tartalom viszont gyakran több sávon oszlik el, az első sáv pedig gyakran csak az intenzitásbeli különbségeket mutatja (4.2.4 fejezet). Transzformált adatsávok helyett ezért minden vizsgált osztály esetében az átlagspektrumtól való eltérést számítottam ki a mintapontokra, az eredményt pedig radiánban mérve prezentáltam (16. ábra).



16. ábra: A vizsgált osztályok homogenitása, spektrális szögeltérés (rad) alapján

A módszer kiugróan magas (0,2 radián fölötti) átlagértéket ad a vizes (v), illetve a részben vízzel borított felszíneknek, mint a nedves talajok (I), mivel a víz reflektancia görbéje általában az abszcissa tengelyt követi, így az egészen kis eltérések is jelentősen megváltoztathatják a spektrális térben mért szöget. A többi osztály homogenitásának becslésére viszont alkalmasnak bizonyult eljárásom, mely mindenekelőtt a természetes növényfajok egy részét (amerikai kőris - A, nemesnyár - J, fehér fűz - K, szürke nyár - M) és a tetőcserepeket (F) jellemezte magas, 0,06 radiánnál nagyobb átlagos szögeltérési értékekkel (4. melléklet).

4.3 Irányított képosztályozások összehasonlítása

4.3.1 Az osztályozások előkészületei

Összehasonlító elemzésemhez a hiperspektrális felvétel nyugati felét, egy 2730 méter hosszú, 640 méter széles pásztát használtam fel (2. melléklet), melyet nagyobb részt természetes és mezőgazdasági vegetáció, kisebb részt talajfoltok borítanak. A vizsgált mintaterület sajátossága, hogy azon jelentős területeket foglaltak el különböző özönnövényfajok állományai, melyek veszélyeztetik az élőhelyek őshonos társulásait és növelik az árvízi kockázatot. A növényzet felmérését a következő fejezetben mutatom be, részletesen kitérve az említett, komplex ökológiai problémára is.

Dolgozatom elemzéseihez egységes szín- és betűkódokat használtam (10. ábra), táblázatukat az egyes osztályozások eredménytérképeihez is mellékeltem. A könnyebb értelmezhetőség érdekében a fekete színt minden esetben az osztályozatlan képelemeknek tartottam fenn. Jelen fejezet osztályozott képeit teljes terjedelmükben a 2. mellékletben mutatom be, az egyébként minden esetben északnak tájolt térképek itt elforgatva láthatók.

Osztályozási eredményeimet kereszt-validációval (cross validation) ellenőriztem vissza, a célom az volt, hogy kiszűrjem a tanító- és tesztpontok kiosztásából adódó esetleges eltéréseket, aránytalanságokat. Az eredeti pontkiosztás mellé még 5 darab tanító-tesztelő ponthalmaz párt generáltam egy MatLab programkód segítségével, az egyes pixelek kiosztása a két halmaz (más néven fold) valamelyikébe véletlenszerűen történt, az arányok ugyanakkor megegyeztek az eredeti 3:1 felosztással, azaz az osztályonként kijelölt kb. 60 db képelem háromnegyede került a tanítópontok közé, negyedét pedig az irányított osztályozások kiértékelésére használtam. A kézzel leválogatott, nagyobb térbeli heterogenitást mutató halmazok rendszerint pontosabb eredménytérképeket adtak, mint az automatikusan generált foldokkal végrehajtott klasszifikációk.

4.3.2 A terepi vegetáció felmérés és ökológiai háttere

Az invazív és egyéb özönnövények terjedésének vizsgálata már évtizedek óta a botanikusok, a természetvédelemmel foglalkozó szakemberek, valamint távérzékeléses környezeti elemzések népszerű és döntéshozók számára is különösen fontos kutatási témái közé tartozik. Az idegen eredetű növényfajok megjelenése elsősorban a távoli földrészek között az újkorban létesülő kereskedelmi kapcsolatokhoz köthető, gyors alkalmazkodásuk és terjeszkedésük pedig a közelebbi rokonok hiányával magyarázható (Rejmánek 1996). A 20. század végén felértékelődött növényi invázió kutatása és nyomonkövetése, amit a témában megjelenő tudományos közlemények számának növekedése is jelez (Pysek et al. 1995, Botta-Dukát 2004).

Az amerikai kőris (*Fraxinus pennsylvanica*) az olajfafélék (*Oleaceae*) egyik, Magyarországon az adventív növények közé sorolt fafaj. Őshazája az Amerikai Egyesült Államok északkeleti, hűvösebb területein található, ez magyarázza kései (májusi) lombfakadását és a viszonylag korai lombhullást is, amely főleg a red ash néven is ismert északi változatára jellemző. A hazai élőhelyek közül a melegebb éghajlatú, vizekben gazdag síkságokat, folyóvölgyeket kedveli, bár lependékes termését nem a víz, hanem jellemzően szél szállítja. Észak-Amerikában kedvelt alapanyagnak számít, Magyarországon viszont könnyen helyettesíthető más fafajokkal, melyek növekedése számottevően gyorsabb is az első évek után. Mindezek miatt a hazai természetvédelem igyekszik visszaszorítani az amerikai kőrist (Csiszár & Bartha 2004).

A gyalogakác, tudományos nevén *Amorpha fruticosa* egy észak-amerikai eredetű cserjefaj, mely a pillangósvirágúak (*Fabaceae*) családjába tartozik. Európába a 18. század első felében került, magyarországi előfordulásáról 1907-ből van először adatunk (Priszter 1997). Mivel az időszakos vízelöntést kedveli, a hazai élőhelyek közül a folyó menti hullámterek galériaerdeiben, elsősorban a telepített nemesnyarasak és a fűz-nyár ligeterdők cserjeszintjében fordul elő, valamint vízelvezető csatornák mentén találhatjuk meg fejlettebb állományait. Mindebben közrejátszik, hogy termése a vízzel terjed, az álló és pangó vizek nyújtotta környezeti feltételek ugyanakkor kevésbé ideálisak a növény számára. Természetvédelmi szempontból a gyalogakác egyértelműen több kárt okoz, mint amennyi hasznót hajt, elnyomja a gyeperdők és a cserjeszint őshonos fajait, akadályozza a természetes erdőfelújítást és növeli az árvízi kockázatot is (Szigetvári & Tóth 2004).

Mintaterületem természetvédelmi kezelője, a Körös-Maros Nemzeti Park Igazgatóság feladatai közé tartozik az özönnövények, mindenekelőtt a hullámterek élőhelyein elterjedt gyalogakác visszaszorítása is. Tevékenységük magában foglalja a megelőzést (kaszálassal és legeltetéssel), a vegyszeres és a mechanikai védekezést, valamint a fertőzött területek erdősítését is. A folyók mentén gyorsan terjeszkedő gyalogakác esetén fontos a folyamatos kezelés, akár évi két alkalommal is, hiszen a kifejlett cserjék a második, augusztusi kaszálás után általában már nem tudnak regenerálódni (Bota 2004).

Terepi felméréseimet már 2010 tavaszán megkezdtem, előbb a Tápai-rét nyugati részén, majd a Tisza hullámterében, a térképezést pedig 2013-ig folytattam, hogy az irányított képosztályozásokhoz minél több és nagyobb kiterjedésű, heterogén tanítóterületet alakíthassak ki. Az invazív növényfajok közül többek között az amerikai kőris állományát sikerült a hullámtérben azonosítanom, amely elsősorban a fűz-nyár galériaerdők környékén jelenik meg, részben elfoglalva a természetes vegetáció élőhelyét. Nagy kiterjedésben találtam a szintén az özönnövények között számon tartott gyalogakác cserjéket, ezek terméseinek szállítódása már jóval nagyobb mértékben kötődik a vizekhez, a nagyobb áradásokhoz, így a hullámtéren kívül a mentett oldali csatornák mentén is számítani kell megjelenésére. Terepi térképezését megkönnyíti, hogy terméshullása akár a következő év nyaráig is eltolódhat, így lombhullás után messziről felismerhetők a visszamaradt fűrtök (17. ábra).



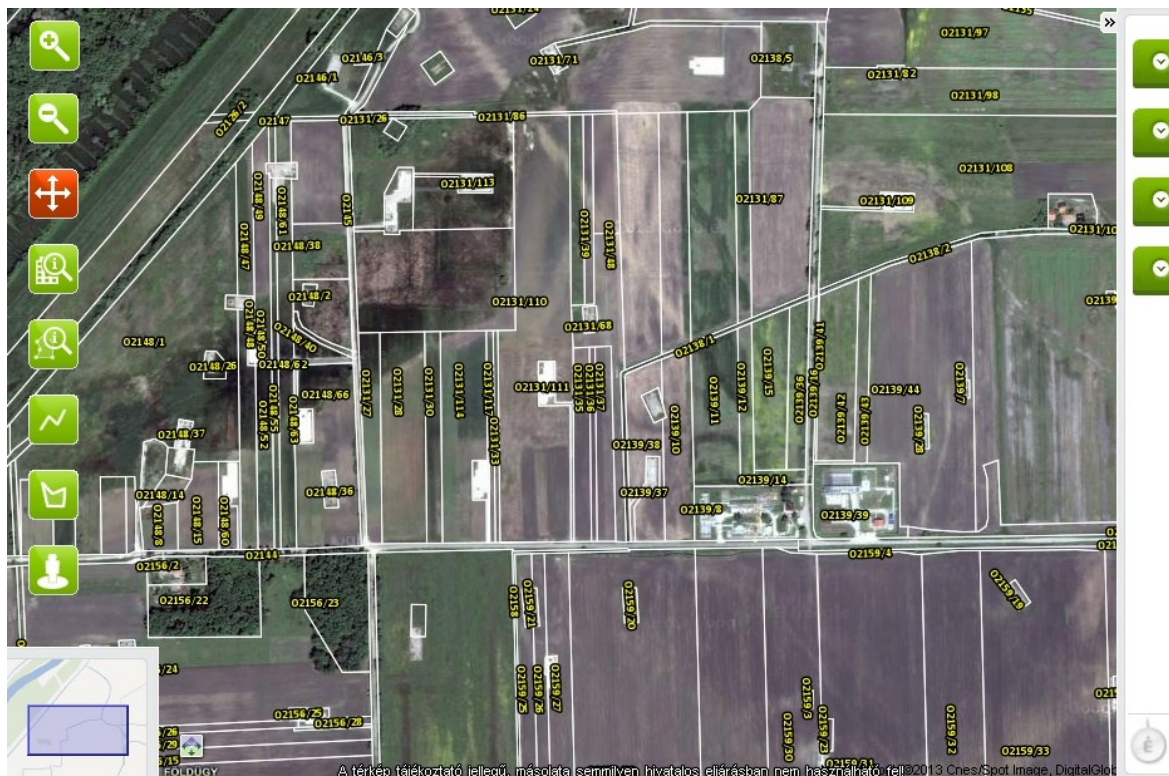
17. ábra: Gyalogakác a telepített nemesnyaras erdő cserjeszintjében (saját felvétel)

Mintaterületem hullámtérre eső részén, az árvízvédelmi töltés víz felőli oldalán a gátépítések idején kiásott, úgynevezett kubikgödrök találhatók. A kubikgödrök környékén ruderalis-félruderalis növényzet telepedett meg, az áradások alkalmával feltöltődő mélyedések vegetációja pedig nagyfokú változékonyságot mutat, a vízborítástól függően akár évről évre más fajok dominálhatják az élőhelyet.

A borításértékeket kb. 10x10 méteres mintavételi egységekre becsültem, amit GPS adatok alapján a légi felvétel térképére vezettem, ezen kívül a nagyobb lombkoronájú, illetve a képosztályozásokban tanítóterületnek használt pixelek esetén egyenként, adott esetben pixelenként határoztam meg a fajt. A vegetáció felméréséhez felhasználtam a MÉTA program növényterképezési adatbázisát is.

4.3.3 Mezőgazdasági adatgyűjtés

A szántóföldi növények tanítóterületeinek kijelöléséhez szükséges információkat a gazdálkodóktól szereztem meg. Első lépésként ehhez az egyes parcellákat kellett azonosítanom, amit az online elérhető Csongrád megyei kataszter segítségével végeztem el. Az adatbázist a Csongrád Megyei Önkormányzat Területi Információs Rendszerének részeként Balog Attila és munkatársai állították össze (18. ábra).



18. ábra: Csongrád megye online kataszteri térképe

A Szegedi Járási Földhivatal közreműködésével ezután felvettem a kapcsolatot a földterületek tulajdonosaival. A begyűjtött adatokat ugyanakkor a légi felvételek alapján is leellenőriztem: a spektrális információk alapján megállapítható, hogy két, az adatszolgáltatók szerint azonos művelésű parcellán valóban ugyanaz a szántóföldi növényfaj található-e. Az összehasonlítást főkomponens-analízissel transzformált képi adatokon végeztem.

4.3.4 Felszínborítás térképezése spektrális szögeltérési osztályozással (SAM)

A spektrális szögeltérésen alapuló osztályozás (Spectral Angle Mapping, röviden SAM) az egyik legelterjedtebb, hiperspektrális felvételeken alkalmazott klasszifikációs módszer, kifejlesztése Kruse és társai nevéhez fűződik (Kruse et al. 1993). Az eljárás alapelve, hogy az egyes spektrális csatornákat adatdimenzióként, a felvételezett reflektancia görbét pedig ebben a multi-dimenzionális térben mint vektort értelmezi (Lillesand et al. 2004). Az osztályozás során a vizsgált spektrumokat egy, az adott osztályra meghatározott referencia spektrummal vetjük össze, a vektorok közti szögeltérést pedig a trigonometriai összefüggések alapján számolhatjuk ki. A szögeltérés értékei értelemszerűen 0 és $\pi/2$ radián között változhatnak, az osztályozás alsó küszöbértékét is ezen a tartományon belül kell meghúzni.

A szögeltérési vizsgálat nem veszi figyelembe a spektrális térben értelmezett vektorok hosszát, ezért kiválóan alkalmas a topográfiai különbségek kiszűrésére (Schowengerdt 2007), az azonos felszíntípushoz tartozó, de eltérő fényintenzitású képelemek összegyűjtésére. Néhány alkalmazás esetén, ahol az intenzitás mértéke fontos minőségi különbségeket jelez, az SAM osztályozást nem célszerű használni, ha azonban az egyes osztályok jól elkülönülnek a spektrális térben, akkor önmagában a szögeltérési vizsgálat is elfogadható pontosságot eredményez (Richards & Jia 2006).

A szögeltérésen alapuló osztályozás egyik legfőbb gyengesége, hogy nehéz meghúzni az osztályozási küszöbértéket, hiszen a spektrális keveredés mértékével általában nemlineáris arányban változik a spektrális szög (van der Meer 2004). Az SAM módszer pontossága ráadásul – hasonlóan számos más képosztályozási eljáráshoz – jelentős csökkenést mutathat, ha a felszíntípusok elkülönítése szempontjából irreleváns spektrális sávok is szerepelnek a vizsgálatban.

A Spectral Angle Mapping (SAM) klasszifikáció mind elméleti hátterét, mind a számításokhoz szükséges időt és hardverkapacitást tekintve a legegyszerűbb algoritmusok közé tartozik. Az eljárás az osztályonként kijelölt mintaterületek spektrumgörbéit átlagolja, majd minden egyes képelemre kiszámolja radiánban a vektorként értelmezett spektrumgörbe és az egyes átlagspektrumok radiánban vett szögeltérését. Mivel a spektrális szögek számításához nincs szükség a képi értékek kovariancia mátrixára, ezért, hasonlóan más,

nonparametrikus módszerekhez, nem szükséges a spektrális sávok számánál több mintapontot kijelölni, ez a tulajdonsága pedig népszerűvé teszi a hiperspektrális felbontású, de gyakran szűk mintaterületre korlátozódó légi felvételek kiértékelésében.

18 osztályt használtam fel az SAM osztályozásban, a mesterséges felszínek közül az utak alapanyagát képező aszfaltot, az olaj- és gázkutak betonplatformját, valamint a piros tetőcserepeket, továbbá a nyílt vízfelszíneket, a talajt, valamint a területen zajló csővezeték építés során megbolygatott talajokat. A mezőgazdasági növények közül a kukoricát, a lucernát, a fokhagymát és a cukorrépát, míg a természetes vegetációval borított felszínekről a nemesnyarast, a szürke nyarat, a fehér fűzet, a tölgyet, a gyalogakácot, az amerikai kőrist, a vízben álló növényzetet és a gyeper osztályát vontam be a térképezésbe.

A képelemek osztályozását egyik esetben sem kötöttem valószínűségi küszöbértékhez, hogy a tesztpontokra kiosztott címkék alapján számszerűsíthessem a felszíntípusok közötti félreosztályozásokat, majd az eredményeket összevethessem a szeparabilitási és a homogenitási statisztikákkal. Az SAM osztályozással 63,6%-os teljes pontosságot, illetve 0,6146-es kappa indexet értem el az általam kijelölt ellenőrző pontokon (5. táblázat). A távérzékeléses vizsgálatoknál szokatlanul alacsony pontosság oka, hogy bár a módszer nem veszi figyelembe az intenzitásbeli különbségeket, a nagyfelbontású adaton egyszerű spektrumátlag alapján nehezebben határozhatók le az egyes osztályok.

Keresztvalidációs halmaz	Teljes pontosság	Kappa index
Fold 0	63,6 %	0,6146
Fold 1	65,44 %	0,634
Fold 2	61,4 %	0,5912
Fold 3	65,44 %	0,634
Fold 4	65,44 %	0,634
Fold 5	65,44 %	0,634

5. táblázat: A szögeltérési (SAM) osztályozás pontosság-értékei

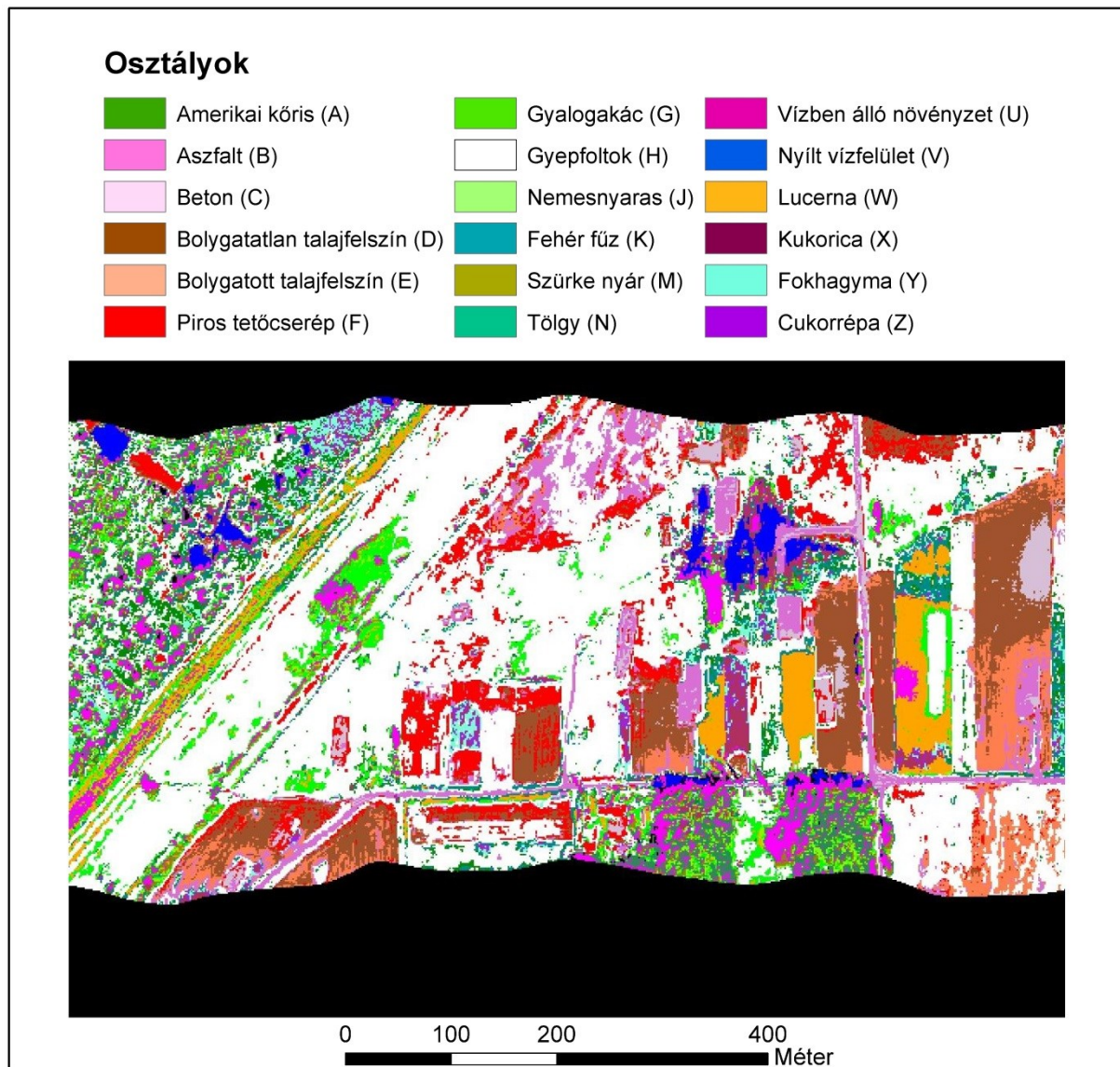
A tévesztési mátrixban (6. táblázat, jelmagyarázata a 19. ábrán) egyes növényfajok osztályozási megbízhatósága (User's Accuracy, U. A.) 50 %-nál is kisebbnek adódott, a lucerna (W), a tetőcserép (F) és a talajfelszínek (D, E) esetében viszont kifejezetten magas találati arányról beszélhetünk, amit egyedi spektrális tulajdonságaikkal, például jellegzetes abszorpciós sávokkal magyarázhatunk.

A spektrumátlag számításán alapuló eljárás előnyeként ugyanakkor megemlíthetjük, hogy a keresztvalidációs halmazok eredményei között elhanyagolható a különbség (5. táblázat), azaz az egyedi minták eltérései kevésbé számítanak, mint például a machine learning algoritmusoknál.

	A	B	C	D	E	F	G	H	J	K	M	N	U	V	W	X	Y	Z	U. A.
A	6							1		2			1						67 %
B		15	5																75 %
C			8																1
D				12	2														86 %
E				3	13														81 %
F			2			14													88 %
G							9	3	4			2	4						41 %
H	3						2	11		2									61 %
J							3		2		5	3	2						13 %
K	2									4									67 %
M									5	2	6	3							38 %
N	1						1		2		1	4							44 %
U									2		3	4	9						50 %
V														17					1
W							1								15				94 %
X	2									5						12	1	3	52 %
Y																	9	5	64 %
Z																3	5	7	47 %
P.A .	43 %	1	53 %	80 %	87 %	1	56 %	73 %	13 %	27 %	40 %	25 %	60 %	1	1	80 %	60 %	47 %	64 %

6. táblázat: Az SAM osztályozás tévesztési mátrixa

Az eredménytérképen a gyepfoltok és a tetőcserép osztályának erős dominanciája figyelhető meg a mezőgazdasági és a talajfelszíneken is (19. ábra), a szerény pontossági eredmények ellenére azonban viszonylag jól elkülönülöníthető néhány parcella (kukorica, lucerna), továbbá sikerült azonosítani a vizeket és a hozzájuk hasonló árnyékos foltokat, valamint a csatornák mentén terjeszkedő gyalogakác bokrokat.



19. ábra: Az SAM módszerrel előállított felszínborítási térkép

4.3.5 Paraméteres osztályozási módszerek

Az irányított osztályozási eljárásokat statisztikai megközelítéseik alapján két nagy csoportba oszthatjuk. Azokat, melyek feltételezik az adatok gaussi eloszlását, statisztikai, vagy paraméteres módszereknek nevezzük, a többi algoritmust pedig a nem-paraméteres vagy neurális osztályozások csoportjába soroljuk (Mather 2004).

A paraméteres osztályozások jelentősen nagyobb számú tanítópontot igényelnek, mint a nem-paraméteres módszerek, hiszen az eljárásokban alkalmazott statisztikai értékek, például a kovariancia mátrix meghatározását is jelentősen befolyásolja a minták számossága. Egyes szerzők a bemeneti adatrétegek, azaz a spektrális sávok számának 30-szorosánál húzzák meg az osztályonként szükséges tanítópontok mennyiségét (Mather 2004), mások sávonként 10 és 100 között határozzák meg azt (Lillesand et al. 2004). Mindez komoly kihívást jelent számos távérzékeléses kutatásban, még úgy is, hogy bizonyos esetekben lehetséges az összes osztály tanítóinak bevonása a statisztikai paraméterek kiszámításába (Cortijo & Pérez de la Blanca 1999).

A statisztikai jellegű megközelítésekhez akkor érdemes ragaszkodni, ha a minták azonosítása nem teljesen megbízható és ez jelentősen rontaná egy nem-paraméteres osztályozás pontosságát. A paraméteres eljárások ráadásul az osztályozáshoz szükséges statisztikai információ túl elméleti hibabecslést is készíthetnek (Schowengerdt 2007). Az elmúlt évtized kutatásaiban azonban a neurális alapú megközelítések kerülnek előtérbe, mivel nem támasztanak követelményeket az adatok gyakoriság eloszlásával szemben (Tso & Mather 2009).

4.3.6 Térképezés Maximum likelihood módszerrel

A hiperspektrális felvétel osztályozását parametrikus módszerrel is elvégeztem, a célra az AISA adatok városi alkalmazásában elért eredmények alapján (Tobak et al. 2013) a Maximum likelihood algoritmust választottam. A statisztikai megközelítésű eljárások jellemzője, hogy a tanítóterületeken végzett előszámításaikhoz a spektrális csatornák számánál több mintapont szükséges. Amennyiben tehát nem sikerül elegendő számú képelemet azonosítani az irányított osztályozáshoz, úgy a spektrális sávok számát kell csökkenteni, amire kézenfekvő megoldás a főkomponens analízis, illetve annak továbbfejlesztett változata, az MNF transzformáció.

A Maximum likelihood osztályozás bemeneti adatának az első 11 PCA sávot választottam, melyek a sajátérték több mint 99%-át tartalmazzák. A 6-11. főkomponens sávok már jelentős zajt is tartalmaznak, mégis célszerű ezeket felhasználni, ha a zaj mellett még érdemi képi információt is őriznek.

A tesztpontokon a magas pontossági eredményeknek megfelelően viszonylag kevés félreosztályozás történt, nagyobb átfedés csupán a kétféle talajfelszín (D és E), a nemesnyár-fehér fűz (J és K) és a fokhagyma-cukorrépa (Y és Z) képelemei között jelentkezik (7. táblázat).

Az osztályozás kappa értéke a manuálisan összeállított tanítóterületekkel 0,9416 lett (8. táblázat), a teljes pontosság 94,49 %, a validációs halmazok számtani átlagát vizsgálva pedig 0,933-at, illetve 93,69 %-ot kapunk. Az egyes fold-ok eredményei között számottevő különbség nem tapasztalható, az Overall Accuracy értékek szórása 1,3 százalékpontos.

	A	B	C	D	E	F	G	H	J	K	M	N	U	V	W	X	Y	Z	U.A.
A	14												1						93%
B		15																	1
C			15																1
D				11	1														92%
E				4	14														78%
F						14													1
G							16												1
H								15											1
J									13	2									87%
K									2	13		1							81%
M											15								1
N												15							1
U													14						1
V														17					1
W															15				1
X																15			1
Y																	13	2	87%
Z																	2	13	87%
P.A.	1	1	1	73%	93%	1	1	1	87%	87%	1	94%	93%	1	1	1	87%	87%	94%

7. táblázat: A Maximum likelihood osztályozás tévesztési mátrixa

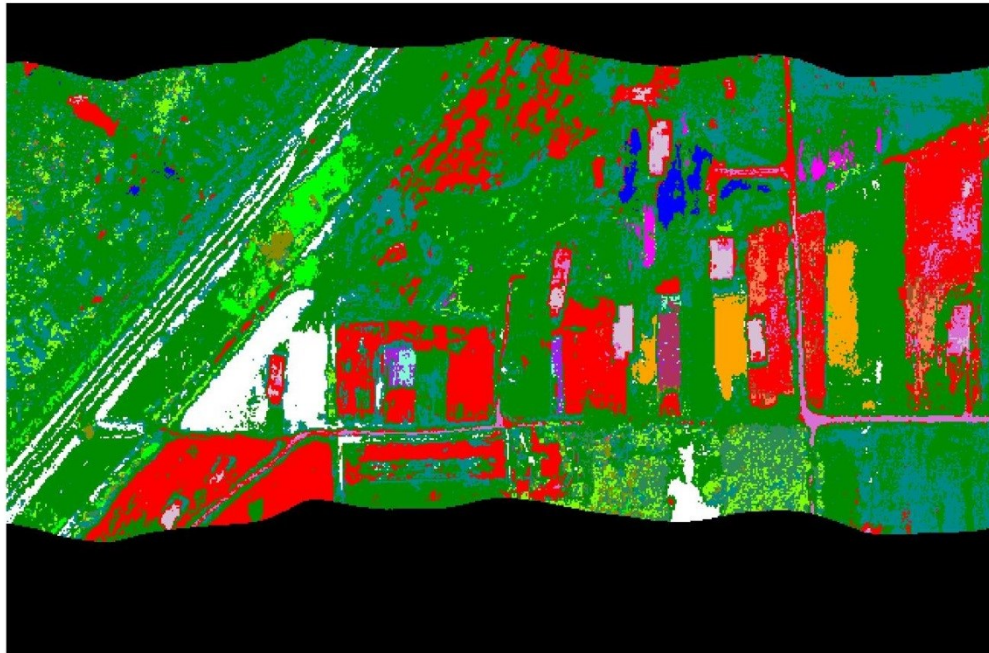
Keresztvalidációs halmaz	Teljes pontosság	Kappa index
Fold 0	94,49 %	0,9416
Fold 1	95,96 %	0,9572
Fold 2	92,65 %	0,9221
Fold 3	93,38 %	0,9299
Fold 4	92,65 %	0,9221
Fold 5	93,01 %	0,926
Számtani átlag	93,69 %	0,933

8. táblázat: A Maximum likelihood osztályozás pontosság értékei

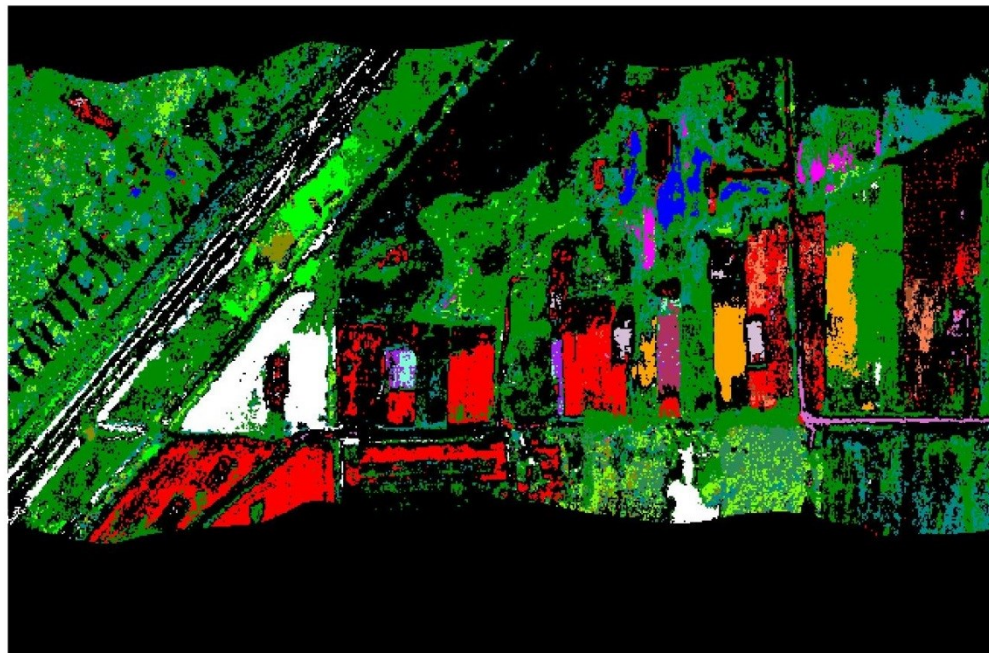
A Maximum likelihood felszínborítási térképén az látható, hogy a képelemek többségét három kevésbé homogén osztály (amerikai kőris, fehér fűz, tetőcserép) címkéi dominálják (20. ábra), a terület nagy része pedig még 40%-os valószínűségi küszöböt alkalmazva is félreosztályozódott, mindez részben a PC sávok használatával, részben a paraméteres algoritmus sajátosságaival függ össze, mely csak a kis számban kijelölt mintapontok szűkebb környezetére volt képes optimalizálni becslési függvényét.

Osztályok

 Amerikai kőris (A)	 Gyalogakác (G)	 Vízben álló növényzet (U)
 Aszfalt (B)	 Gyepfoltok (H)	 Nyílt vízfelület (V)
 Beton (C)	 Nemesnyaras (J)	 Lucerna (W)
 Bolygatatlan talajfelszín (D)	 Fehér fűz (K)	 Kukorica (X)
 Bolygatott talajfelszín (E)	 Szürke nyár (M)	 Fokhagyma (Y)
 Piros tetőcserép (F)	 Tölgy (N)	 Cukorrépa (Z)



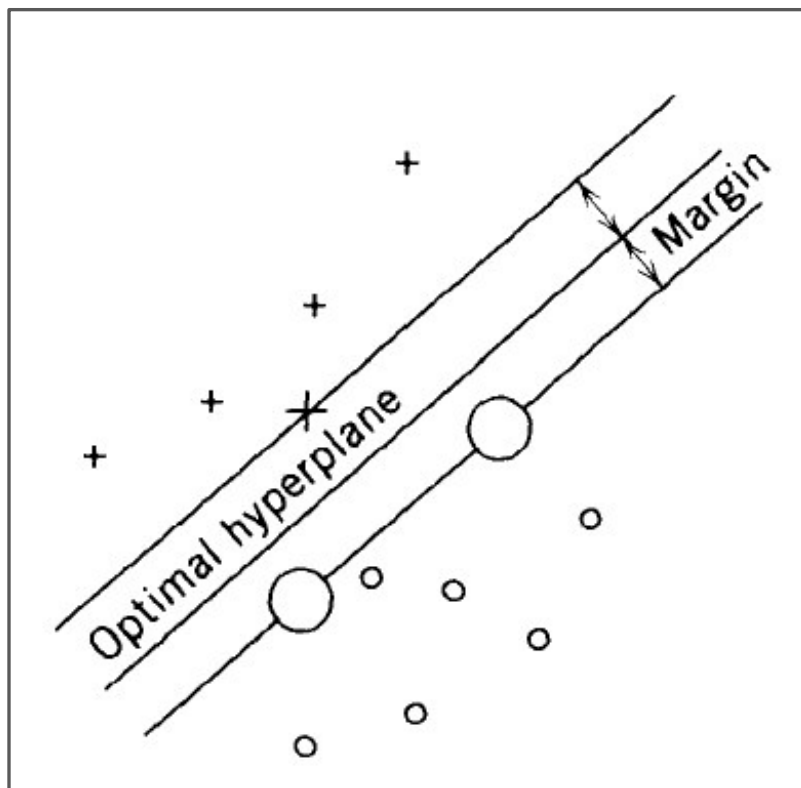
0 100 200 400
Méter



20. ábra: A Maximum likelihood osztályozás eredménye (kivágat), a középső képen valószínűségi küszöb nélkül, az alsó képen 40%-os küszöbvel

4.3.7 Képosztályozás Support Vector Machine alkalmazásával

A Support Vector Machine (SVM) napjainkban az egyik legelterjedtebb machine learning algoritmus a képfeldolgozás területén, noha a statisztikai tanulás elméletébe illesztett módszer már a hetvenes években ismert volt (Vapnik 1998), távérzékelt adatok osztályozására elsőként Gualtieri és Cromp (1999) alkalmazta. Az SVM eljárás úgy minimalizálja az osztályozási hibát, hogy a tanítópontoktól lehető legnagyobb távolságban („margin”) jelöli ki az osztályok elválasztását szolgáló úgynevezett optimális hipersíkot (21. ábra).



21. ábra: Az optimális hipersík elhelyezkedése (Vapnik 1998)

Büntető paraméterekkel (penalty parameter) lehet szabályozni a modell illeszkedését a tanítópontokhoz, a lineárisan szét nem választható adatok osztályozása pedig a megadott kernel függvény alapján, magasabb adatdimenzionalitás mellett történik (Vapnik 1998). A bias-variance tradeoff elmélet szerint modellválasztásnál a strukturális hibák mellett a varianciából adódó túlilleszkedést is figyelembe kell venni, ami SVM osztályozások esetén ugyanolyan jelentőségű lehet, mint a pontatlanul kijelölt tanítóterületekhez köthető tévesztések (Cawley & Talbot 2010).

Az osztályozás paramétereit úgynevezett grid search eljárással optimalizáltam, majd miután a legmagasabb pontosságot elértem (97,06%), ezeket a beállításokat követve alkalmaztam az öt további keresztvalidációs halmazon is (9. táblázat). Teljes pontosságuk és kappa értékük, bár elmarad a fold 0-val jelzett eredeti pontkiosztásától, minden esetben 94%, illetve 0,93 fölött van. A 18 osztály közül 12 esetében semmilyen félreosztályozás nem történt (10. táblázat), a terepi információ és az osztályozás oldalán sem, szerényebb mértékű tévesztés is csupán a két talajfelszín osztály (D és E), két mezőgazdasági növény (fokhagyma és a lucerna), valamint az amerikai kőris és a gyepfoltok viszonylatában történt. A magas találati pontosság részben az algoritmus túlilleszkedésével (overfitting) függ össze, azaz az eljárás csak a betanításra kijelölt mintapontokat és szűkebb környezetüket képes megbízhatóan azonosítani, az ezeken kívül eső, de azonos felszínborítottságú képelemek jelentős arányban félreosztályozódnak.

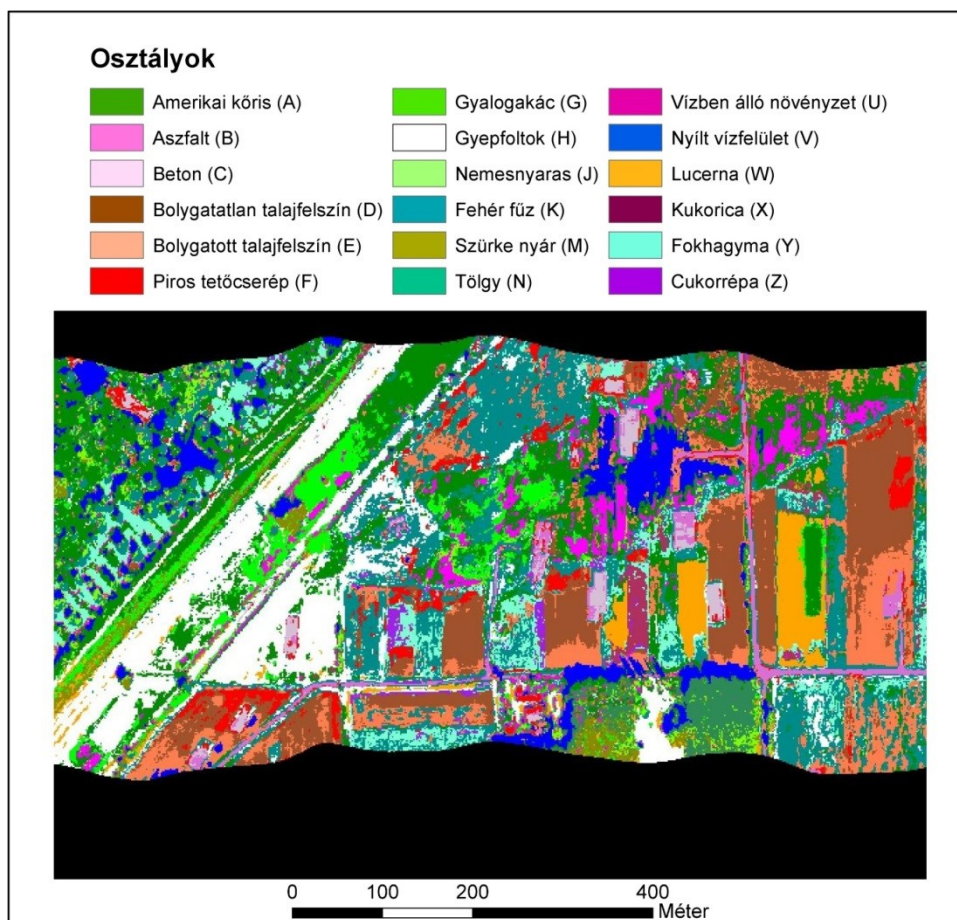
Keresztvalidációs halmaz	Teljes pontosság	Kappa index
Fold 0	97,06 %	0,9689
Fold 1	95,96 %	0,9572
Fold 2	94,85 %	0,9455
Fold 3	96,32 %	0,9611
Fold 4	95,22 %	0,9494
Fold 5	94,12 %	0,9377

9. táblázat: Az SVM osztályozás pontosság értékei

Az SVM eredménytérképén viszonylag nagy kiterjedésben találunk a tetőcserepekhez, vagy a fehér fűz osztályához rendelt képelemeket a mintaterület középső és keleti, talajfelszínekkkel és mezőgazdasági növényekkel borított részein (22. ábra), a képosztályozás számára kihívást jelentett a szűk tanítóterületek alapján azonosítani ezeket a foltokat, amiben közrejátszhatott a képi zaj és a területi autokorreláció is. A fehér fűz és a gyep, továbbá a tetőcserepek és a talajfelszín közötti átfedések ugyanakkor hangsúlyosan jelentek meg a két főkomponens sávon lefuttatott Jeffries-Matusita szeparabilitási vizsgálatban is (3. táblázat).

	A	B	C	D	E	F	G	H	J	K	M	N	U	V	W	X	Y	Z	U.A.
A	13																		1
B		15																	1
C			15																1
D				13	2														87%
E				2	13														87%
F						14													1
G							16												1
H	1							15											94%
J									15										1
K										15									1
M											15								1
N												16							1
U													15						1
V														17					1
W															15				1
X																15			1
Y																	14	2	88%
Z																	1	13	93%
P.A.	93%	1	1	87%	87%	1	1	1	1	1	1	1	1	1	1	1	93%	87%	97%

10. táblázat: Az SVM osztályozás tévesztési mátrixa



22. ábra: Az SVM osztályozás eredménye (kivágat)

4.3.8 SOM-hibrid osztályozás és klaszterezés

Az SOM (Self-Organizing Map) kifejezés a mesterséges neurális hálózatok (ANN) egy különleges válfaját takarja. Az általános ANN formula így adható meg (Tso & Mather 2009):

$$x_j = \sum_i a_i w_{ji} - \theta_j$$

ahol x a neuronra érkező teljes kimeneti jel, a az aktivitási érték, w az egyéni súly, θ pedig az eltolási tényező.

A Kohonen által a 80-as évek elején kifejlesztett, és a kilencvenes évek óta a távérzékelésben is alkalmazott eljárás (Kohonen 1997, Chiuderi & Capellini 1996), amellelt, hogy kiosztja az adatleíró prototípusokat, indexálja is azokat. Az alacsony dimenzionalitású SOM rács, melybe a prototípusokat rendezzük, már alkalmas az egészen nagy méretű és magas dimenzióval jellemezhető sokaságok áttekinthető térképezésére. Ezen a kimeneti rétegen az egyes elemek hasonlóságaik alapján vannak rendszerezve, a Kohonen-féle eljárás tehát úgy tárja fel és ábrázolja a sokaság sűrűség-eloszlását, az adatklaszterek topológiáját, hogy mindeközben nem csökkenti a prototípus vektorok dimenzionalitását (Merényi et al. 2009).

Az SOM tanulási folyamata formálisan a következő algoritmussal írható le (Tso & Mather 2009):

$$d_j^2 = \sum_i^k (x_i^n - w_{ji}^n)^2$$

ahol d^2 a bemeneti és a kimeneti neuronok közötti euklideszi távolság négyzete, x az adott neuron bemeneti jele, w az i -edik neuronról a j -edikre érkező jel súlyértéke, k pedig a bemeneti adat dimenzionalitása (Tso & Mather 2009). A tanulási folyamat során tehát nem csupán a bemeneti adat nemlineáris projekciója történik meg, a neurális hálózat elemei is egyfajta tömbbe rendeződnek a hozzájuk rendelt prototípusok (súlyvektorok) topográfiai ismérvei alapján. Az egyes lépésekben azonosított neuronok ugyanis nem csupán maguk számára tárolják el a bemeneti információt, hanem meghatározott sugarú körben szomszédait is aktiválják, azaz a tanulás során rendező, finomító hatást a szomszédsági információ használata vált ki. A hatótávolság meghatározása az SOM vizsgálat egy különösen érzékeny pontja. A túl nagy távolsági érték elmoshatja a különbséget, míg a szűkre szabott

aktiválási környezetben a topológia-térkép mozaikossá válik, ahol a klaszterek között nem feltétlenül alakul ki folyamatos átmenet (Kohonen 1997).

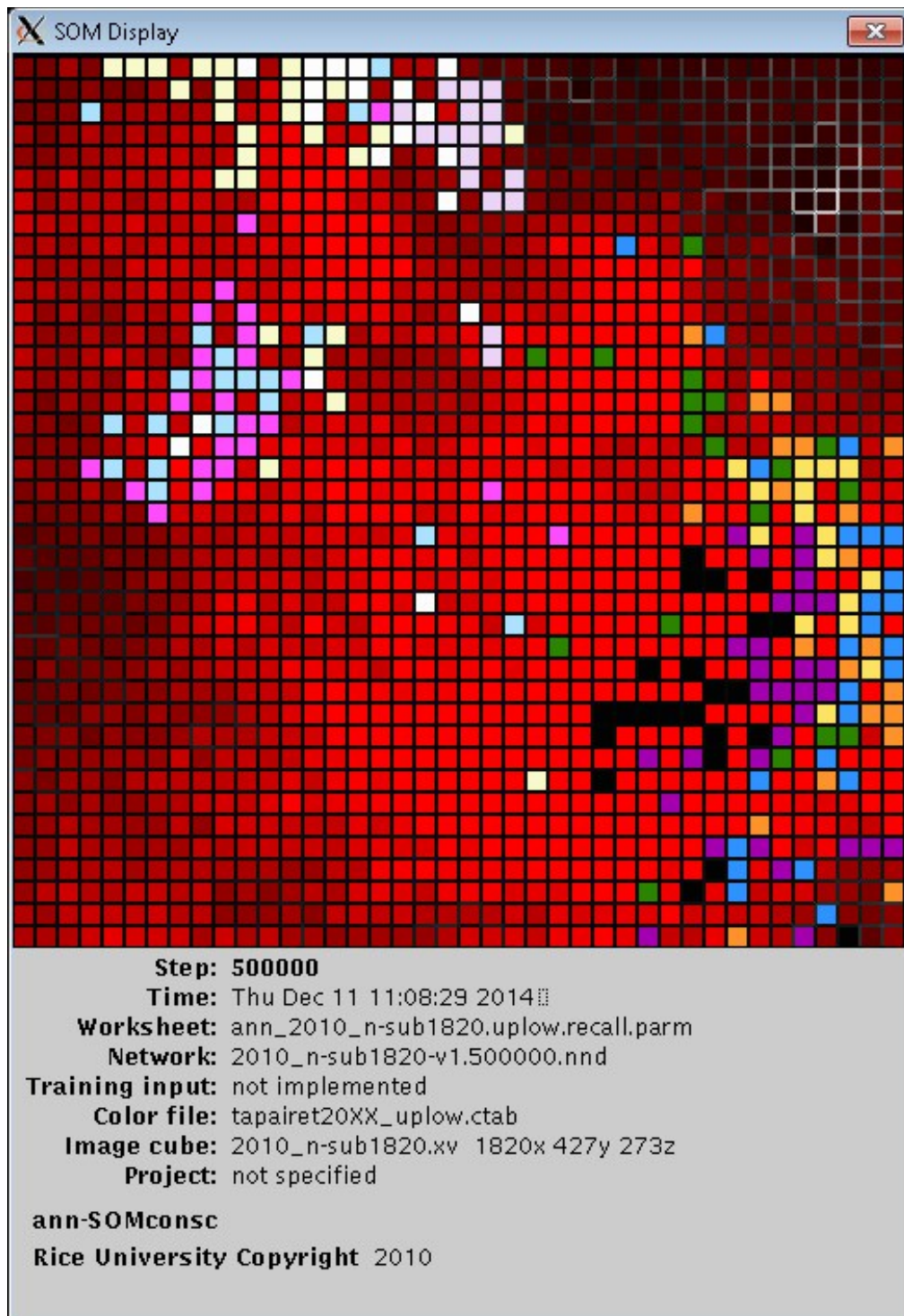
A klaszterek elkülönítéséhez nem feltétlenül kell pontosan meghatározni az adat topológiai összefüggéseit, viszont a sűrűség eloszlásához már alaposabb tanítási folyamat szükséges. A topográfiai leképezés helyességének vizsgálatára, illetve a pontosság számszerűsítésére leteznek eljárások, mint például a Topographic Function, ami az SOM-ben jelentkező átlagos topológiai eltérést mutatja meg (Villmann et al. 1997). Az anomáliákat (topology violation) két típusra oszthatjuk: ha a bemeneti adatra kifeszített Voronoi-poligonok szomszédos középpontjai a leképezés során eltávolodnak egymástól, forward topology violationról beszélünk, míg backward topology violation esetén az SOM rácsban olyan elemek kerülnek egymás mellé, melyek Voronoi-középpontjai nem szomszédosak a bemeneti adaton (Merényi et al. 2009).

Az SOM osztályozást a houstoni Rice Egyetem statisztikai tanszékén végeztem, az itt alkalmazott Khoros képfeldolgozási-programozási környezet ugyanis – Merényi Erzsébet, Patrick O’Driscoll és kollégáik sokéves kutatói-fejlesztői munkájának köszönhetően – alkalmas a különösen nagyméretű képi állományok feldolgozására is.

A zajos sávoktól megtisztított, adathiányos szegélyeinél kimaszkolt felvételeket importáltam a Khoros fordított bájtsorrendű rendszerébe. A Rice Egyetemen kifejlesztett CLSTAT modulban kiértékeltem a spektrumgörbéket, majd a statisztikai információk alapján elvégeztem a reflektancia adatok normalizációját, elősegítve ezzel a neuronok gyorsabb tanulását. Az osztályozás irányított szakaszához szükséges volt importálni a tanító- és tesztpontok koordinátáit is, melyeket úgynevezett include file-okba rendeztem, valamint összeállítottam az osztályok szín- és betűkódjait tartalmazó colortable (.ctab) listát.

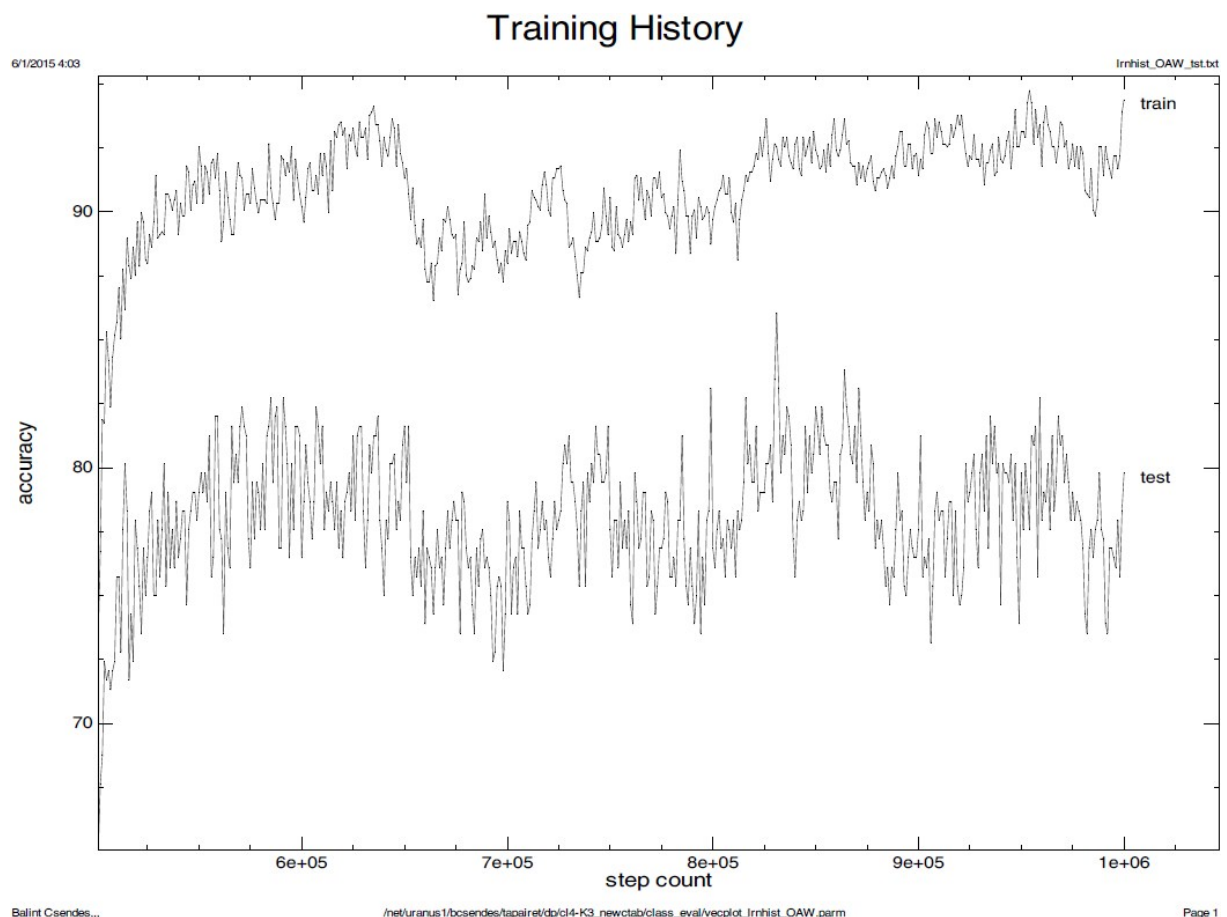
Az osztályozás első lépése az irányítatlan tanulási eljárás, melynek során egy kétdimenziós mátrixban is megjeleníthető topológiai hálót rajzol fel az algoritmus a bemeneti adatok eloszlása alapján (Schaale & Furrer 1995). Inputnak a Khoros rendszerbe importált (.kdf kiterjesztésű), zajos sávoktól megtisztított, normalizált képet használtam. A programot korábban orvosi képfeldolgozási kutatásokban is használták (agyi fMRI felvételek klaszterezése), az ezekben felmerülő nehézségek, adatfeldolgozási korlátok elkerülésére szűkítettem a vizsgált mintaterületet, levágva annak az elemzés szempontjából érdektelen, keleti szélét. Az így kapott, $1820 \times 427 = 777140$ képelemen végzett tanítás eredményét két alkalommal ellenőriztem a topológiai térkép előhívásával (recall), az 500 ezredik lépést

követően, amikor még biztosan nem került minden pixel kiértékelésre, illetve a 8 milliomodik tanítási esemény után. A félmilliomodik lépésnél kirajzolódó topológiai mátrixot mutatja vörös árnyalatú képén a 23. ábra, ahol a színes cellák jelzik az klaszterek összetartozó elemeit, a jobb felső sarokban látható világosabb és fehér határvonalak pedig a szomszédos klaszterek közötti gyengébb kapcsolatra vagy a kapcsolat hiányára utalnak.



23. ábra: Az SOM klaszterek topológiai mátrixa

Az SOM térképezés irányított fázisában az azonosított klaszterek címkézésére (labeling) kerül sor, a megjelölésen túl azonban a súlyok változtatásával azok elhelyezkedésén is módosítani lehet, igazodva a tanító osztályokhoz. A súlyértékek korrekciójának mértékét szabályozó δ paramétert három lépésben csökkentettem, hogy a tanítási iterációk előrehaladtával egyre finomabban közelítsem a maximális pontosságot. A tanító- és a tesztpontokra számított teljes pontossági százaléértékek (Overall Accuracy) változását a 24. ábra mutatja az 500 ezredik lépéstől kezdődően (az irányított tanítást a félmillió iteráció után előhívott klasztertérképen kezdtem meg). A tesztpontok optimuma a 831 ezredik lépésnél található (86,03%), ugyanakkor a tanítóterületekre számított pontosság itt nem kiugró (92,52 %). Az irányított tanítást a többi keresztvalidációs halmazon is elvégeztem, azonos paraméterekkel számolva, kiválasztva a legmegfelelőbb scenáriót (11. táblázat). A pontosság ezeknél elmaradt az eredeti pontkiosztáshoz képest, a további öt halmaz átlaga azonban így is 80% fölött van.



24. ábra: Az SOM irányított tanítása, a tanító- és tesztpontok százalékos pontosságának függvényében

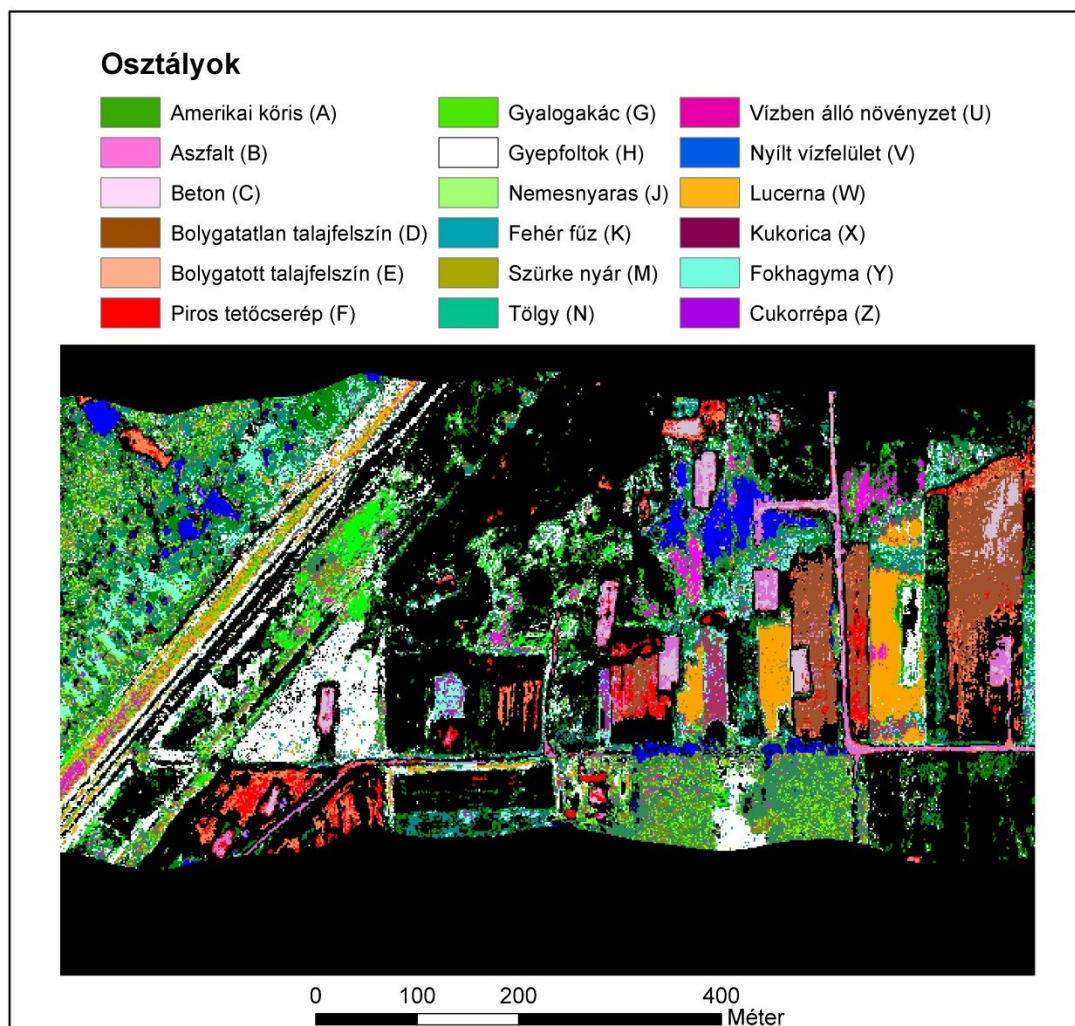
Keresztvalidációs halmaz	Teljes pontosság	Kappa index
Fold 0	86,03 %	0,8521
Fold 1	83,09 %	0,8209
Fold 2	77,57 %	0,7625
Fold 3	83,46 %	0,8248
Fold 4	83,09 %	0,8209
Fold 5	80,88 %	0,7975
Számtani átlag	82,35 %	0,813

11. táblázat: Az SOM-hibrid osztályozás pontosság értékei

	A	B	C	D	E	F	G	H	J	K	M	N	U	V	W	X	Y	Z	U. A.
A	9																		1
B		14						1											93 %
C		1	15																94 %
D				10	1														91 %
E				5	14														74 %
F						14													1
G							13				1								93 %
H							2	14	1										82 %
J									9		1	2							75 %
K	2									14	1	1	1						74 %
M	1						1		3	1	12	1							63 %
N	1								2			12							80 %
U	1												14						93 %
V														17					1
W															15				1
X																13	1		93 %
Y																1	13	3	77 %
Z																1	1	12	86 %
P.A .	64 %	93 %	1	67 %	93 %	1	81 %	93 %	60 %	93 %	80 %	75 %	93 %	1	1	87 %	87 %	80 %	86 %

12. táblázat: Az SOM osztályozás tévesztési mátrixa

A vizsgált osztályok mindegyikén sikerült legalább 60%-os pontosságot elérni (12. táblázat), kiugró tévesztési arányokat az amerikai kőris (A) és a nemesnyár (J) sorában találhatunk. Az SOM algoritmus osztályozatlanul hagyta a tanítóterületek mintáitól spektrális értékükben távolabb eső képelemeket, ezzel, bár bizonyos esetekben romlik a képosztályozás pontossága, nagyban növeli a felszínborítási térkép megbízhatóságát. Eredménytérképét szemlélve megállapítható, hogy az esetenként túlterjeszkedő tetőcserepeket (F), illetve a fokhagymát (Y) leszámítva nem történt nagy kiterjedésű területeket érintő félreosztályozás a kivágaton (25. ábra), több olyan foltot viszont nem vagy csak részben sorolt be valamelyik osztályba, melyeket más algoritmusok felismertek, mint például a töltés mentén húzódó gyep a kivágat nyugati részén, vagy a délnyugati talajfelszínnek.



25. ábra: Az SOM osztályozás eredménye (kivágat)

4.3.9 A felszínborítási térképezések diszkussziója

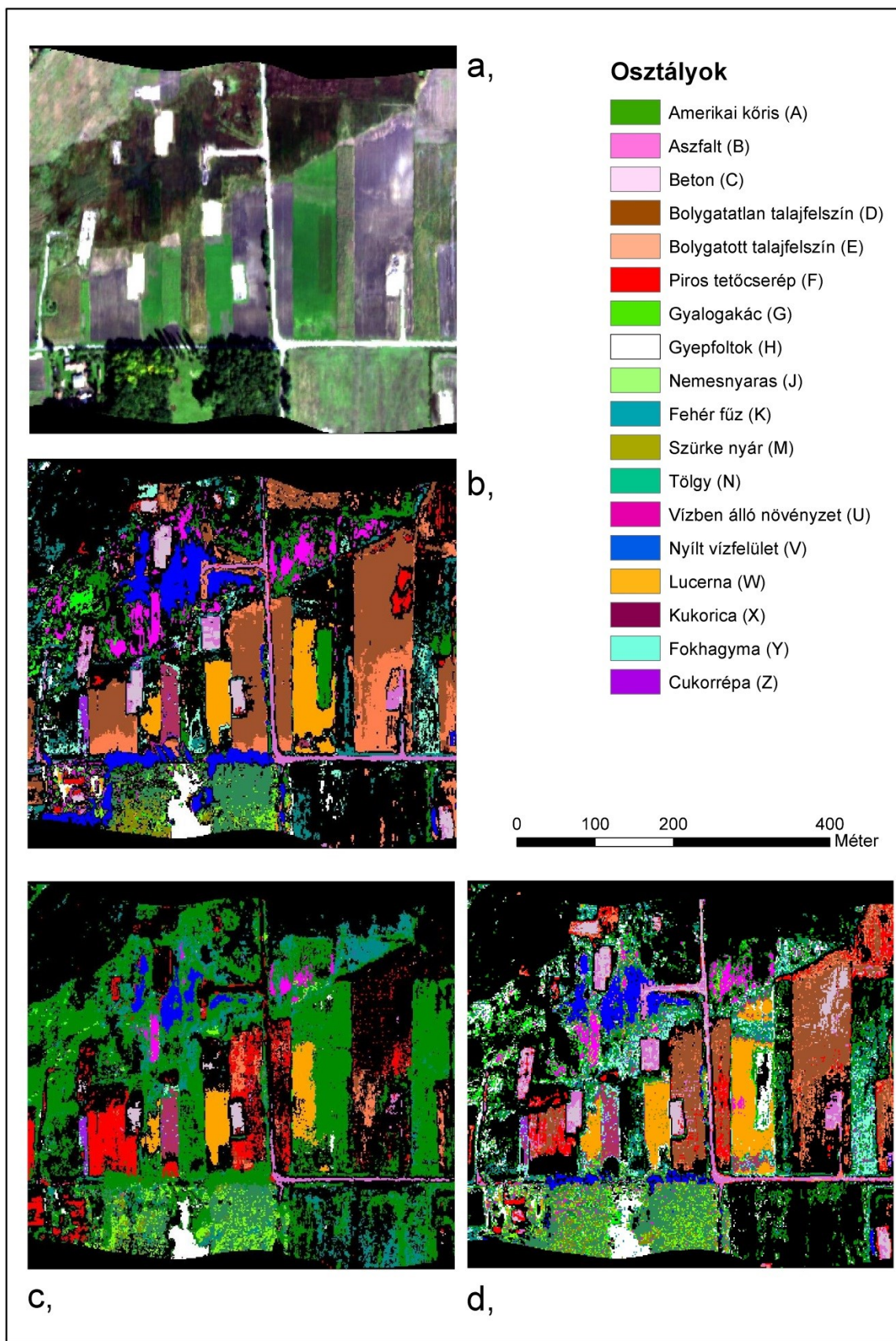
A képosztályozási módszerek pontosságát tárgyaló tudományos közlemények következtetései szerint a Machine Learning algoritmusok megfelelő, az esetek többségében pontosabb és megbízhatóbb alternatívát kínálnak az egyszerűbb statisztikai megközelítésű eljárásokkal szemben (Lary et al. 2015, Merényi et al. 2007a).

Magyarországi mintaterületen végzett képosztályozásaim is alátámasztották a fenti következtetéseket, teljes pontosság szempontjából a Maximum likelihood (94,5%), az SVM (97%) és az SOM hibrid módszer (86%) is lényegesen jobb eredményt adott, mint az átlagspektrumokkal operáló SAM (63,6%).

A Maximum likelihood eljárás átlagosan csupán a mintapontok 6 százalékát osztályozta félre, annak ellenére, hogy bemeneti adatként mindösszesen csak az első 11 főkomponens sávot használtam. Összehasonlítva a többi térképkiívággal, 40%-os valószínűségi küszöbérték mellett is jól kivehető egyes osztályok túlilleszkedésből fakadó dominanciája (26. ábra). Más tényezők mellett ez rávilágít a feature reduction okozta információ veszteségre és a csak kis mértékben zajos hiperspektrális adatsávok szerepére a spektrális diszkriminációban (Howell et al. 1994, Moon et al. 1995, Merényi et al. 2014).

SVM osztályozással az esetek többségében hatékonyabban sikerült a felszín típusokat felismerni és elkülöníteni, mint más klasszifikációs algoritmusokkal (Pal & Mather 2005, Pal & Foody 2010, Cortes & Vapnik 1995), előnyei magasabb dimenzionalitású adathoz még hangsúlyosabbak (Huang et al. 2002). A módszer korlátai modellválasztásnál, a paraméterek meghatározásánál jelentkeznek (Cawley & Talbot 2010), a neurális hálózatok ezzel szemben, a nagyobb léptékű tanulási eljárásuknak köszönhetően, jobban adaptálhatóak a nem minden részletében ismert, összetettebb felszínborítású mintaterületeken.

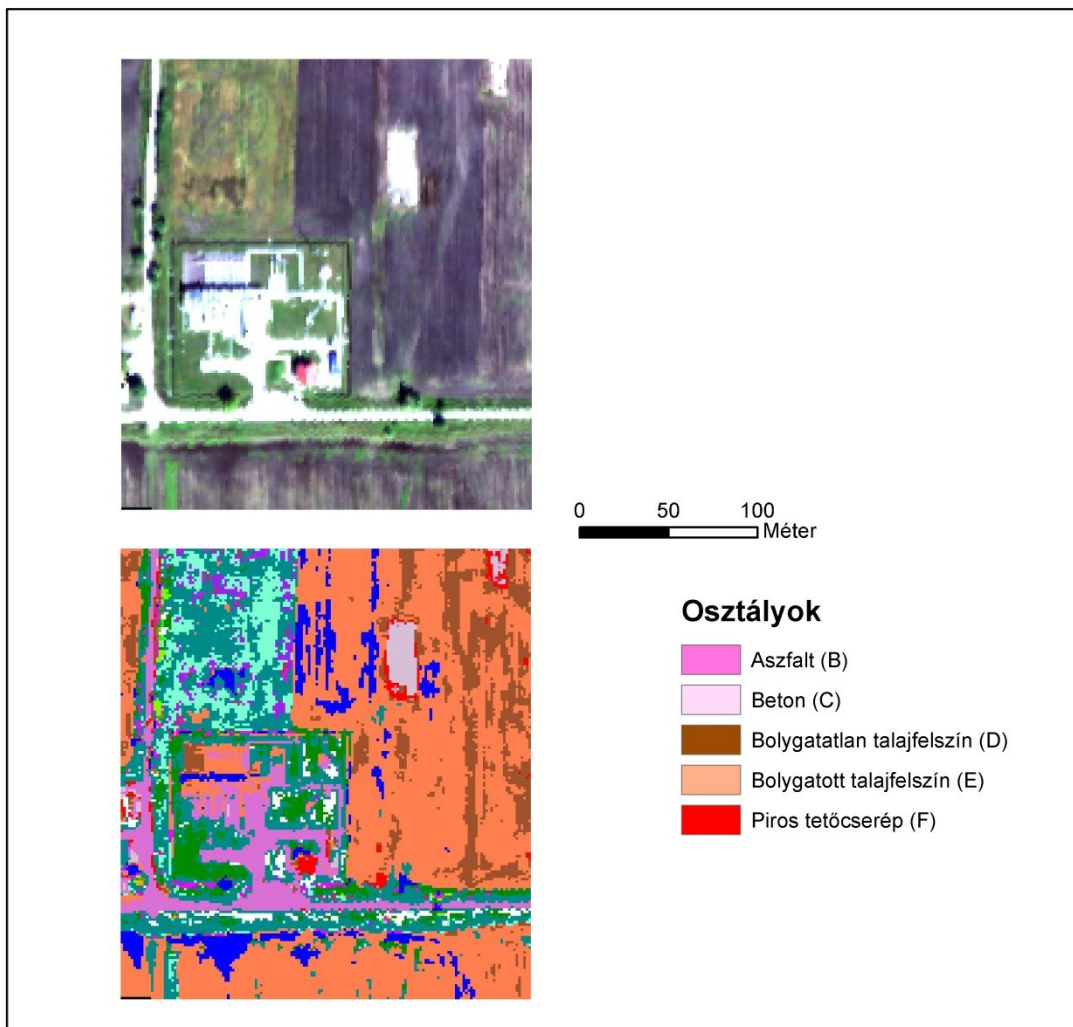
Erőforrás használat szempontjából a nem-paraméteres, irányított osztályozások bizonyultak hatékonyabbnak, az SAM és az SVM módszerek ugyanis nem igényeltek hosszadalmas adattranszformációkat, sem iterációs műveleteket.



26. ábra: Osztályozott képrészletek összehasonlítása: a, valószínűségi képkivágat b, az SVM térképe c, a Maximum likelihood térképe d, az SOM osztályozás eredménye

4.3.10 Mesterséges felszínek térképezése

A Tápai-réten azonosított három mesterséges felszíntípus, a piros tetőcserepek, az aszfalt utak és a szénhidrogén bányászathoz kialakított betonplatformok elkülönítése magas pontossággal történt valamennyi osztályozási módszerrel, a szeparabilitási vizsgálatok viszont jelentős spektrális átfedéseket mutattak a tanítóterületeik között, különösen a két főkomponensen végzett Jeffries-Matusita-féle számítások esetében (3. táblázat).



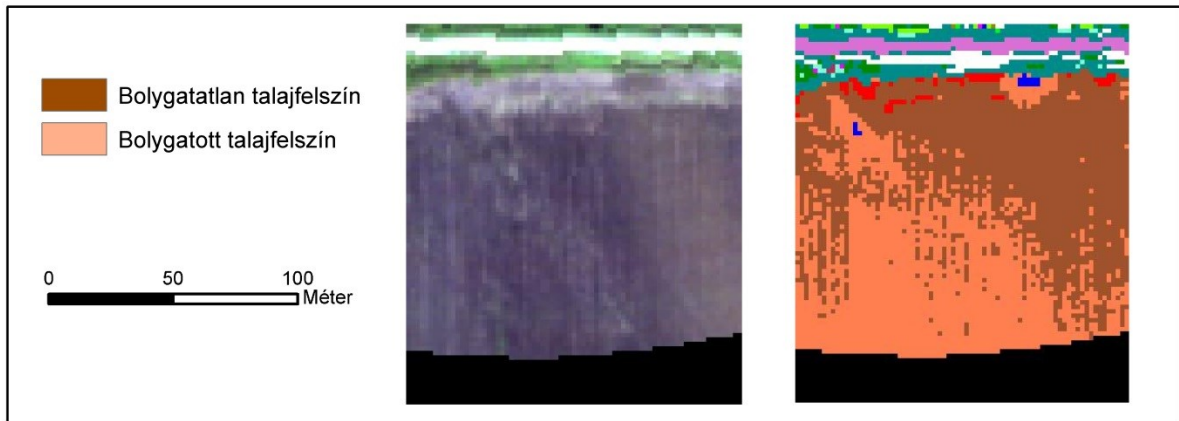
27. ábra: Mesterséges felszínek az SVM eredménytérképen

Az elkülöníthetőségi statisztika következtetéseit az eredménytérképek igazolták vissza, melyeken a nyílt talajok képelemeinek jelentős részét mind a négy eljárás (SAM, ML, SVM, SOM) a három nevezett osztály valamelyikébe sorolta (26. ábra), miközben az ellenőrzőpontok adatain a félreosztályozásaik egyáltalán nem jelentek meg. A 27. ábra a magasabb pontosságú SVM módszerrel készült felszínborítási térkép egy képrészletét mutatja, az itt található telephely és környezetének aszfalt, beton és tetőcserép borítású felszínei jól elkülönülnek, a beton és a tetőcserép osztályok között azonban kisebb mértékű félreosztályozások látszódnak (27. ábra).

4.3.11 Átmozgatott talajok osztályozása

A tájtérképezési vizsgálatba a csővezeték-építés során megbolygatott talajfelszíneket is bevontam. Az átmozgatott talajokról és a kontrollterületekről, melyeket bolygatatlan felszínnek neveztem el, 15-15 db mintapontot vettem fel a Tápai-rét déli részén, egy röviddel a felvételezés előtt felszántott mezőgazdasági parcellán. A Transformed Divergence szeparabilitási index enyhe (1,93), a Jeffries-Matusita módszer szignifikáns spektrális átfedést (1,6) mutatott a két osztály között, melyeket a spektrális szögeltérések alapján szerkesztett box plot diagram szerint nagyfokú homogenitás jellemez (16. ábra).

A két felszíntípust a képosztályozásokkal sem sikerült maradéktalanul elkülöníteni, az SVM algoritmus, mely a teljes adaton 97%-os teljes pontosságot mutatott, az ellenőrzőpontok 2-2 elemét összecserélte, így mindkét osztály pontossága és megbízhatósága (user's accuracy) is 87 %-osnak adódott. Az SVM-mel készült tematikus térképen, bár felfedezhető az építkezés északnyugat-délkeleti irányú nyomvonala, a két felszíntípus határvonala mégis elmosódott és bizonytalan, a félreosztályozott területek könnyen felismerhetők (28. ábra). Az eredmények tükrében megállapítható, hogy képosztályozásaim az átmozgatott talajfelszínnek megbízható azonosításásra nem alkalmasak, a viszonylag magas találati pontosság pedig a jól elkülönített tanítóterületeknek és az algoritmusok túlilleszkedésének köszönhető.

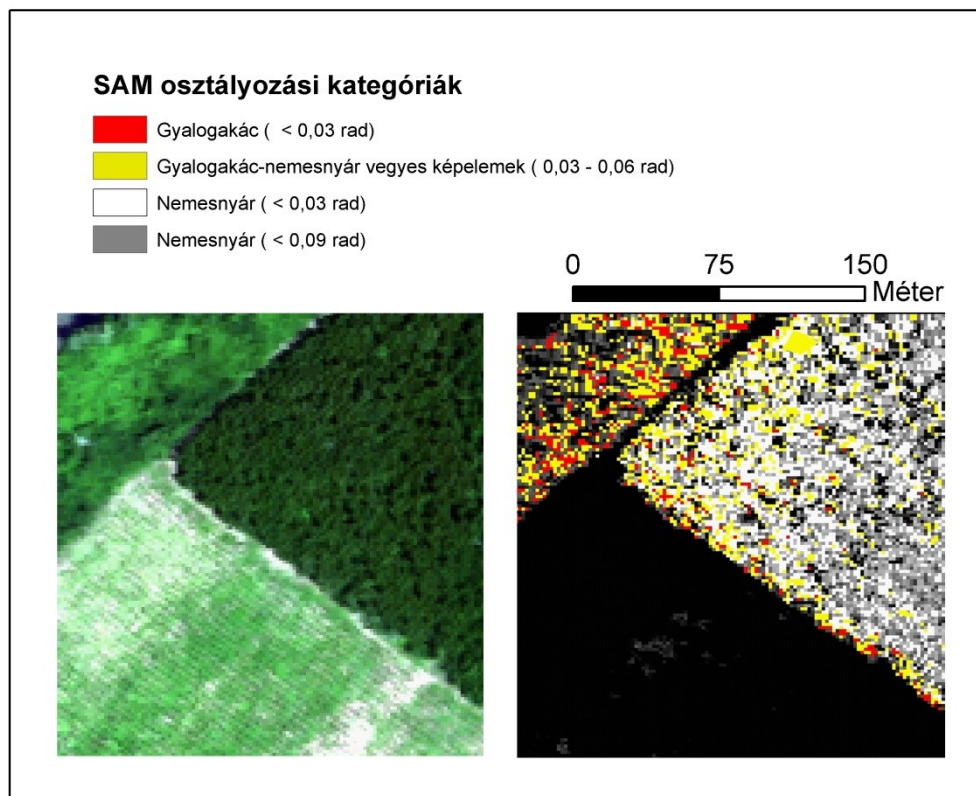


28. ábra: A talajfelszínek térképkivágata, SVM módszerrel osztályozva

4.3.12 A lombkorona alatti cserjeszint spektrális vizsgálata

A távérzékelésen alapuló növényterképezésben különleges kihívást jelent a lombkoronaszint alatt található vegetáció azonosítása, osztályozása, azonban nem példátlan a spektrálisan többé-kevésbé kevert pixelek elkülönítése a többi, magasabbra nöövő fajok által dominált képelemtől (Williams & Hunt 2002).

Kutatásomat a tápai-réti mintaterület északi részéről 2009-ben készült AISA hiperspektrális felvételeken (Csendes 2013) végeztem. A szükséges előfeldolgozási lépések után az adaton spektrális szögeltérési vizsgálatot (SAM) hajtottam végre, azért ezt az osztályozási módszert választottam, mert a túlilleszkedés elkerülésére alkalmasabbak a spektrumátlag alapú eljárások (4.3.9 fejezet).



29. ábra: A gyalogakác és a nemesnyár elkülönítése SAM osztályozással

Vizsgálataimhoz egy olyan telepített nemesnyaras erdőfoltot választottam ki a mintaterületemen, ahol a cserjeszintet gyalogakác alkotja, ahogyan a 4.3.2 fejezet 17. ábráján is látható, a gyalogakác bokrok térképezését célzó kutatásaimban ugyanis visszatérő probléma volt, hogy a magasabbra növő fák részben vagy teljesen eltakarták a vizsgált növényeket (Csendes 2012). Terepi felméréseim alapján kijelöltem a két osztály, a két növényfaj tiszta pixeleit, majd a kivágot többi elemére is kiszámítottam a két átlagspektrumtól mérhető spektrális szögeltérést. A 29. ábrán vörös szín jelöli azokat a képelemeket, ahol az eltérés a gyalogakác átlagspektrumától $0,03$ radiánnál kisebb, a fehér foltok a nemesnyáréhoz hasonló reflektanciájú területek jelzik, míg a sárgára színezett felszínek $0,03$ és $0,06$ rad közötti spektrális szögeltérést mutatnak mindkét osztálytól.

Az osztályozott térkép alapján megállapítható, hogy a vizsgált erdőfolt képelemeinek többségén a lombkoronaszintet alkotó nemesnyarasok dominálnak, $10,9\%$ -uk spektrális szögeltérése mindkét növénytípus átlagától számított $0,03 - 0,06$ rad között található, míg a pixelek $1,5$ százalékán a gyalogakác átlagspektrumától $0,03$ radiánnál kisebb eltérés mérhető.

Ez utóbbiakat jellemzően az erdőfolt szélén sikerült azonosítani, néhány esetben azonban a nemesnyaras belsejében is találtam összefüggő gyalogakác foltokat (29. ábra).

Osztályok	Nemesnyár	Kevert képelemek	Gyalogakác	Transformed Divergence index
Nemesnyár	0	0,413	1,717	Nemesnyár
Kevert képelemek	0,408	0	1,361	Kevert képelemek
Gyalogakác	1,431	0,883	0	Gyalogakác
Jeffries- Matusita index	Nemesnyár	Kevert képelemek	Gyalogakác	Osztályok

13. táblázat: A lombkorona- és a cserjeszint osztályainak szeperabilitása

A szögeltérés alapján leválogatott osztályok átfedéseit spektrális szeperabilitási indexekkel vizsgáltam (13. táblázat), a két növényfaj mellett a kevertebb képelemek elkülöníthetőségét kívántam elemezni. A Jeffries-Matusita módszer ezúttal is érzékenyebb volt a Transformed Divergence (TD) indexnél, eredményeik azonban egybevágnak, a gyalogakác és a nemesnyár értékei is jelentős, 1,9 alatti értékű átfedést mutatnak, a legalacsonyabb mutatószámok ugyanakkor várakozásaimnak megfelelően a kevert képelemekenél találhatók. Mindkét számításnál, de különösen a TD esetén tapasztalható kiugróan alacsony szeperabilitás a vegyes pixelek és a nemesnyár mintái között, ez alapján megállapítható, hogy a spektrális szögeltéréseik alapján leválogatott kevert képelemeken is a lombkoronát alkotó nemesnyarak dominálnak. Ennek magyarázata az lehet, ha reflektancia görbéiken megjelenik egy, például a gyalogakác illóolajtartalmához köthető elnyelési sáv, ami a spektrális szöget is módosítja, a transzformált adatsávokra számított intenzitásértékek osztályon belüli eloszlása viszont a másik növényfajára jellemző.

4.4 Belvítérképezés irányítatlan és irányított képosztályozásokkal

4.4.1 A belvív, mint környezeti probléma

A belvív meghatározására több tucat definíció létezik, figyelembe véve a jelenség hidrológiai, geomorfológiai, biológiai, klimatikus és gazdasági vonatkozásait (Rakonczai 2013, Pálfi 2001). A belvív képződésében éppúgy szerepet játszhat az intenzív csapadék vagy a kedvezőtlen vízgazdálkodású talaj, mint az agrogén eredetű talajtömörödés és a mélyen fekvő területek nem megfelelő vízelvezetése. Ha a belvizet, mint környezeti problémát vizsgáljuk, célszerű az ökológiai, illetve közgazdasági megközelítésű definíciókat használni. Ezek alapján belvíznek azt a tartós, lefolyás nélküli vízborítást nevezzük, amely kimutatható károkat okoz (Rakonczai 2013), beleértve a teljesen átnedvesedett talajok okozta vegetációkárosodást is (Rakonczai et al. 2011). Hozzá kell azonban tenni, hogy bizonyos esetekben, amikor például vizes élőhelyek, természetvédelmi oltalom alatt álló területek vízutánpótlásáról van szó, a fenti szempontok alapján a belvízképződés hasznos is lehet (Rakonczai 2013).

A belvív többféleképpen hat a mezőgazdasági területek termőképességére, mely hatások között találunk pozitívokat és negatívakat egyaránt (Szatmári & van Leeuwen 2013). A tartós belvízelöntés kiszorítja az oxigént a talaj felső rétegeiből és ugyanitt jelentős sófelhalmozódást eredményezhet, ami jelentősen visszavetheti a növények fejlődését. A káros hatások között említhetjük még a járműközlekedés akadályozását, valamint a szennyezőanyagok gyorsabb terjedését (Mucsi 2001). A környezeti feltételeknek megfelelően megválasztott, víznek jobban ellenálló növényfajok termesztésével ugyanakkor mérsékelni lehet a károkat, sőt, a belvízképződést előidéző hidrometeorológiai tényezők növelhetik is a terméshozamot. Esetenként a termésmennyiség növekedése figyelhető meg az érintett területek közelében, de azoknál valamivel magasabban fekvő térszíneken, valamint azokon a mezőgazdasági parcellákon is, melyeket előző évben belvízelöntés sújtott. A globális klímaváltozással összefüggésben várhatóan növekedni fog a téli és tavaszi csapadék mennyisége a Kárpát-medencében (Bartholy et al. 2007), ezért a jövőben kiterjedtebb belvízképződéssel kell számolni a késő ősztől kora tavaszi terjedő időszakban (Szatmári &

van Leeuwen 2013), mindamelllett az intenzívebbé váló csapadékhullás is növeli az időszakos vízelöntések kockázatát (Somlyódy et al. 2010). Az elöntési kockázat mérséklésére számos eszközt ismerünk, mint a szivattyúzás, a felszíni csatornázás, dréncsövek lefektetése vagy vízigényes növényfajok telepítése (Szatmári & van Leeuwen 2013). A belvízprobléma hosszú távú megoldását azonban a körütekintő vízgazdálkodás jelentené, ami magában foglalja a víztöbblet visszatartását a későbbi, aszályos időszakokban történő felhasználáshoz (Rakonczai 2013), és amely legalább annyira figyelembe veszi a természetvédelmi szempontokat, mint a mezőgazdasági károk minimalizálását.

4.4.2 Terepi belvízfelmérés

A belvíz egy időszakos jelenség, ezért a térképezése is csak bizonyos, nedvesebb periódusokban lehetséges. A belvízelöntések jelentős időbeli és térbeli változékonyságot mutatnak, egy tájegységen belül akár évszakonként más térszíneken jelenhet meg, ugyanakkor a vegetáció változása és a talajfelszín cserepedése jelzi a gyakori vízborítást, illetve a túlnedvesedést. Szárazabb periódusok idején ez a két indikátor segítette a terepi felmérést, míg nedvesebb időszakokban, mint amilyen a 2010-es év, illetve 2013 első féléve volt, a nyílt vízelöntéseket is térképezni lehetett (30. ábra).



30. ábra: Belvízelöntés a mintaterületen, két nappal a felvételezés előtt

(Készítette: Tobak Zalán, 2010. szeptember 20.)

Az adatrögzítést Garmin Oregon 450 és Garmin eTrex GPS mérőműszerekkel végeztem. A mintaterületeket a GPS adatok és a terepen készített fényképek alapján jelöltem ki, figyelembe véve a légi felvételek térbeli felbontását és az ebből adódó esetleges spektrális keveredést is.

4.4.3 A távérzékeléses belvízkutatások előzményei

A talaj nedvességtartalmának térbeli eltéréseit azonosítani lehet nagyfelbontású távérzékelte adatokon, ezt sikerült support vector regresszióval (Stamenkovic et al. 2013), fajlagos emisszió számításával (Sobrino et al. 2012) és diszkrét wavelet transzformáció esetén (Peng et al. 2013) is bizonyítani. *Tits et al (2013)* mindezt kiegészítették azzal, hogy a száraz és a nedves talajok osztályozásánál, melyet spektrális szétválasztással végeztek, a magasabb reflektanciájú szárazabb felszíneket nagyobb abszolút, viszont alacsonyabb relatív hibaarányjal jellemezték.

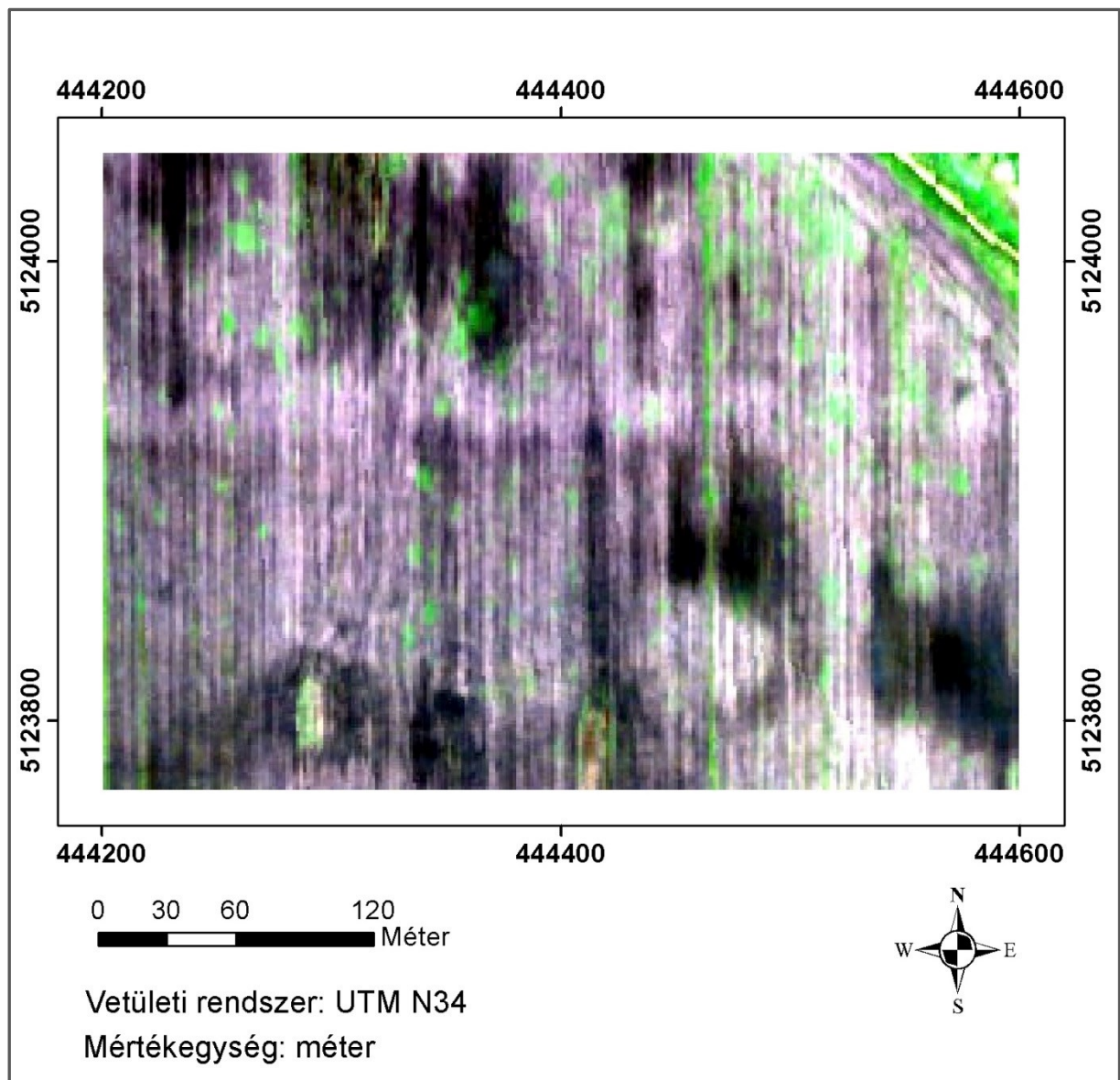
A Szegedi Tudományegyetem geoinformatikai kutatócsoportja a belvízveszélyes területek terepi felmérésével és távérzékeléses vizsgálatával is foglalkozik. Az 5 méteres geometriai felbontású RapidEye műholdképeken irányítatlan, irányított, valamint mesterséges neurális hálózatokon (ANN), illetve spektrális szétválasztáson (SMA) alapuló módszerekkel is térképezték a Szeged környéki belvízfoltokat és egyéb, túlnedvesedett talajfelszíneket (Mucsi & Henits 2011, van Leeuwen et al. 2012). A Maximum likelihood és az ANN-alapú klasszifikációik hasonlóan magas pontosságú (96-99%) eredménytérképet adtak, kevés átfedés csupán a víz- és a talajfelszín osztályai között jelentkezett, az irányítatlan osztályozások alkalmazását ugyanakkor teljesen elvetették, mivel az előállított klaszterek nem estek egybe a vizsgált felszíntípusokkal és megfelelő terepi felmérési információ sem állt a kutatók rendelkezésére a kvantitatív kiértékeléshez (van Leeuwen et al. 2013a, van Leeuwen et al. 2013b).

A spektrális szétválasztási módszerekkel rosszabb felbontású (Landsat TM és ETM+) műholdfelvételeken is készíthető belvízelöntési és vízborítottsági aránytérkép (Mucsi & Henits 2011), akár országrész nagyságú területek érintettségének becslésére (Csendes 2010).

A találati pontosság további javulását eredményezte, amikor a színes infravörös légi felvételek mellett különféle földrajzi adatbázisok térképeit is bevonták a vizsgálatokba, közülük is elsősorban a domborzati viszonyokat leíró LIDAR adatok és a vízelvezető csatornáktól mért távolság felhasználása járult hozzá érdemben az ANN osztályozás eredményességéhez (van Leeuwen et al. 2012).

4.4.4 Hiperspektrális adataim a belvizes területekről

Könnyen azonosítható belvízfoltok a mintaterület keleti szélén voltak megfigyelhetők, ahol a hiperspektrális légi felvételezés idején sem a természetes, sem a mezőgazdasági növények nem takarták el a túlnedvesedett talajfelszínt, ezért képosztályozásomat is ezen a kivágaton hajtottam végre (31. ábra). A terület északnyugati sarkában és a déli részen található nagyobb kiterjedésű vízelöntés, ezen kívül észak-déli irányú barázdák mentén fedezhetők fel sötétebb színű, belvizes pixelek. A képrészlet területe mindössze 11 hektár, amin belül kb. 60-60 db nedves, illetve vízelöntéssel egyáltalán nem érintett talajokat ábrázoló képelem különíthető el egyértelműen, a terület nagy részét nem lehet minden kétséget kizáróan valamely kategóriába besorolni, hiányoznak továbbá a nagyobb mélységű vízfoltok spektrális szempontból tiszta, keveredésmentes pixelei is.



*31. ábra: A belvívelöntéssel érintett mintaterület valószínűségi képkivágata
a hiperspektrális adaton*

4.4.5 Irányított belvízosztályozások

A belvível borított szántóföldeket ábrázoló képkivágaton tanítópontokat vettem fel a jellemző felszíntípusokról, így a nyílt vízborítású területekről, a nedves, ill. a száraz talajfoltokról és a részben vízzel borított növényzetről. A minták segítségével két, egyszerűbb, túlilleszkedésre kevésbé hajlamos, nem-paraméteres osztályozást végeztem el, a

spektrumgörbe értékeit a spektrumátlag alapján felosztó bináris spektrumkódolást (Binary Encoding) és a spektrális szögeltérési vizsgálatot (SAM).

Tanítópontjaim kijelölését nehezítette, hogy a belvízfoltok kiterjedése jelentős időbeli változékonyságot mutatott, továbbá a szűkebb mintaterületen belül sem volt lehetséges a felszíntípusokat pontosan elhatárolni, ezért az osztályozási adatbázisba csak azokat a képelemeket válogattam be, melyek a terepi felmérések és egyéb képi adatforrások alapján is egyértelműen besorolhatók valamelyik belvíztérképezési kategóriába. Bemeneti adatnak ebben az esetben is a zajtól mentes 273 spektrális csatornát használtam. Osztályok közötti átfedéseket a kevésbé érzékeny Transformed Divergence szeparabilitási számítással nem, csupán a Jeffries-Matusita indexszel sikerült kimutatni (14. táblázat). A két főkomponensen elvégzett vizsgálat a nedves és a száraz talajok, valamint a nedves talajfelszínek és a nyílt vizek között mért a teljes elkülönülést jelző 2-es értéknél alacsonyabb, 1,9 körüli szeparabilitást.

Jeffries-Matusita	Vízfelszín	Túlnedvesedett talaj	Száraz talaj	Vízben álló növényzet	
Vízben álló növényzet	2	2	2	0	Vízben álló növényzet
Száraz talaj	2	1,891	0	2	Száraz talaj
Túlnedvesedett talaj	1,897	0	2	2	Túlnedvesedett talaj
Vízfelszín	0	2	2	2	Vízfelszín
	Vízfelszín	Túlnedvesedett talaj	Száraz talaj	Vízben álló növényzet	Transformed Divergence

14. táblázat: A belvíztérképezési osztályok szeparabilitási értékei az első két főkomponens sáv alapján

A bináris spektrumkódolás az egyik legegyszerűbb statisztikai megközelítés, mellyel a spektrumgörbéket osztályozhatjuk. Az egyes reflektancia értékeket a következő módon kategorizáljuk (Richards & Jia 2006, Goetz et al. 1985):

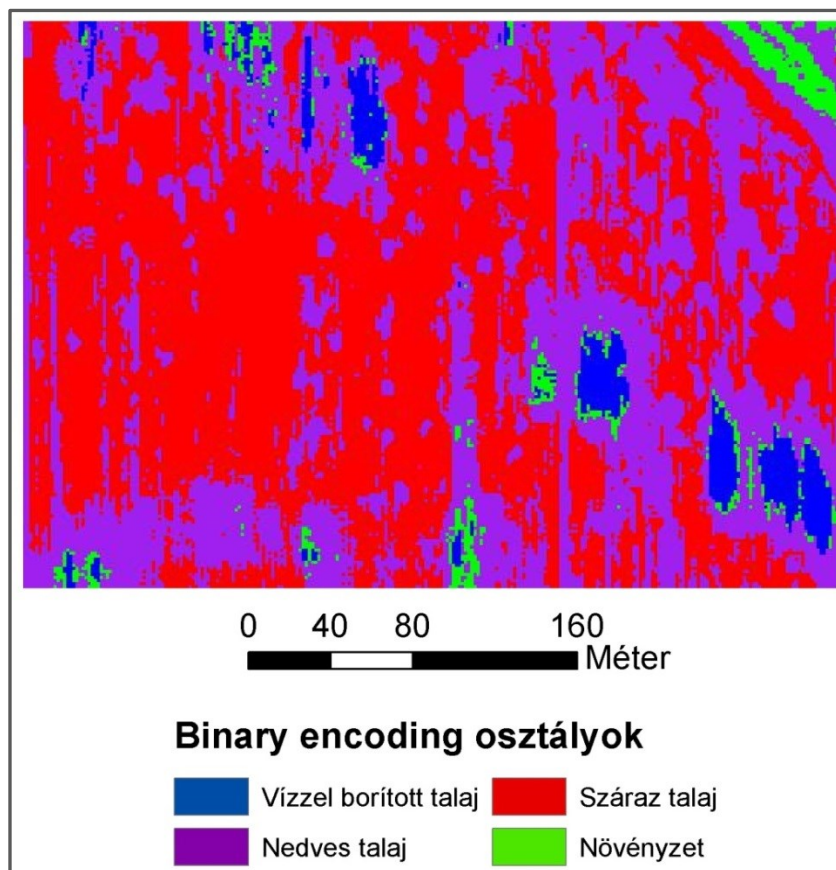
$$h(n) = \begin{cases} 0 & \text{ha } x(n) \leq T \\ 1 & \text{egyébként} \end{cases} \quad n = 1, \dots, N$$

ahol $h(n)$ az n -edik spektrális sávhoz tartozó új kódérték, $x(n)$ a képelem reflektancia arányszáma az adott sávban, T pedig a megválasztott küszöbérték (Richards & Jia 2006).

Az eljárás akkor terjedt el széleskörűen, amikor a számítógépek adatfeldolgozási kapacitása még nem tudott lépést tartani a rohamosan bővülő távérzékelési adatbázisokkal. A leegyszerűsített spektrumkódok tárolása ugyanis jóval kisebb memóriát, a spektrumok összehasonlítása kevesebb időt igényelt, a küszöbérték körültekintő megválasztásával pedig jó pontossággal el lehetett különíteni az elnyelési sávokat (Schowengerdt 2007).

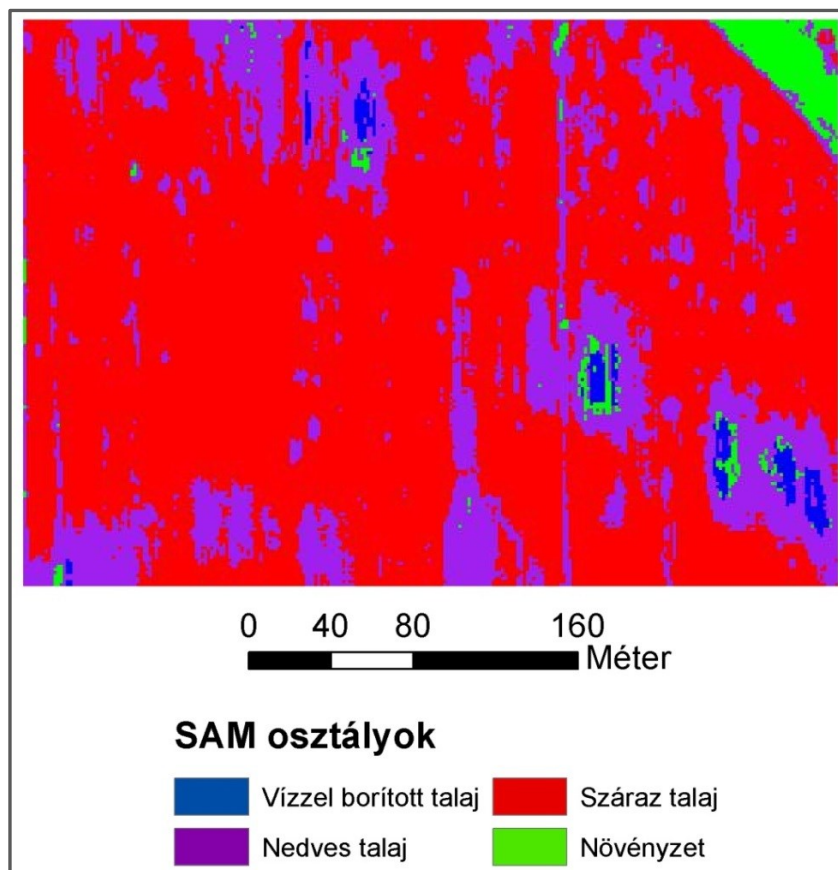
A bináris kódolás továbbfejlesztésére több módszert is kidolgoztak. A besugárzás változékonysága és az atmoszférikus hatások miatt érdemes az adott spektrumgörbe átlagos reflektancia értéke, nem pedig egy, az egész felvételre érvényes küszöbérték alapján végezni a kódolást (Mazer et al. 1988). A számítási- és memóriaigény növekedése árán javíthatjuk úgy is a pontosságot, ha egyszerre több küszöbértéket is alkalmazunk (Jia & Richards 1993). A kódolás alapjául ráadásul nem csupán magát a reflektanciát, hanem például annak deriváltját is használhatjuk, 0 kódértékkel jelezve a negatív, 1-gyel pedig a pozitív függvénymeredekséget (Schowengerdt 2007).

Binary Encoding osztályozásnál a pontosság 90, az SAM esetében pedig 84 százalékosnak adódott a száraz és a túlnedvesedett talajfelszínek viszonylatában. A magasnak mondható találati értékeket a rendelkezésre álló tanító és ellenőrző pontok kis száma magyarázza, az erre a célra felhasználható képelemek szűk körében ugyanis egymásra nagyon hasonlító, izolált foltjaikban területi autokorrelációt mutató pixeleket találunk.



32. ábra: Binary encoding osztályozás eredménye a belvizes területen

Az osztályozási térképeken azonosíthatók a valószínűségi képkivágaton (31. ábra) is könnyen felismerhető, nedvesebb felszínek, a Binary Encoding esetén kiterjedtebbek ezek a foltok (32. ábra), míg az SAM képén a száraz talajok dominálnak (33. ábra), a nedves és a száraz talajfelszínek elkülönítése azonban mindkét módszernél bizonytalan, a jellegzetes észak-déli irányú barázdák nem minden esetben rajzolódnak ki. Magyarázata mindennek az lehet, hogy a kis számban leválogatott mintapontok nem fedik le a különböző nedvességtartalmú talajok teljes változatosságát, a finom átmenetek térképezéséhez sokkal részletesebb, terepi mintavételezéssel kiegészített térbeli adatbázisra lenne szükség.



33. ábra: SAM osztályozás eredménye a belvizes területen

4.4.6 Irányítatlan belvízosztályozás ISODATA módszerrel

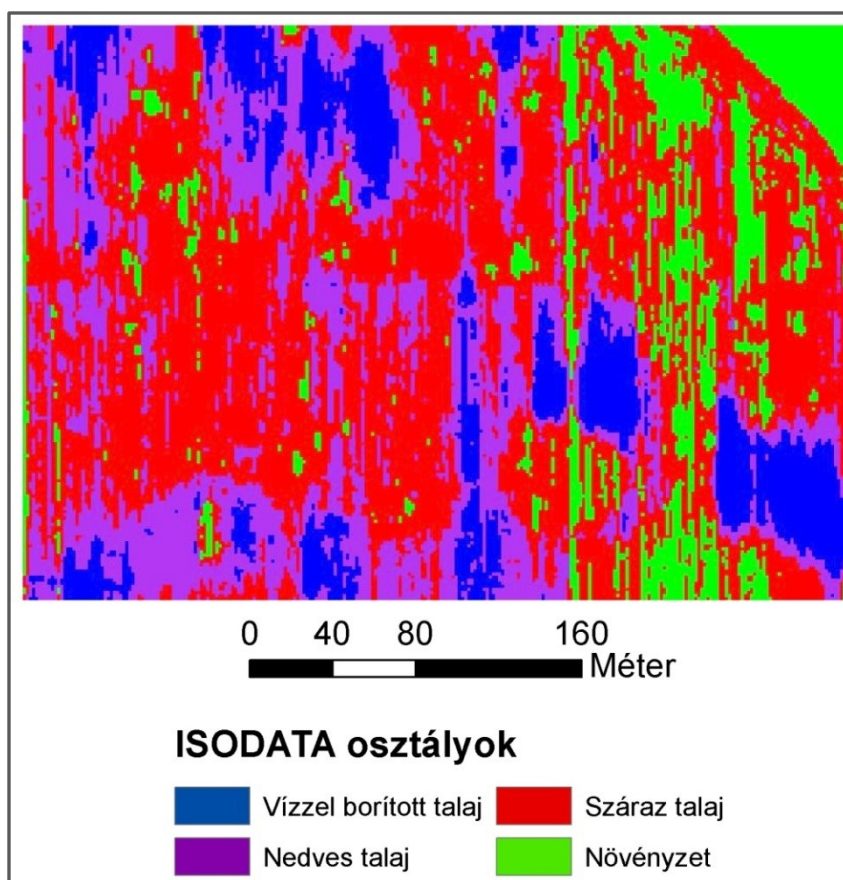
Osztályozásokat előre definiált tanítópontok nélkül is végezhetünk, az irányítatlan klaszterezési eljárások legnagyobb előnye, hogy a mintaterületen előforduló felszíntípusok teljes körű ismerete nélkül is térképezhetővé teszik a hasonló spektrális tulajdonságú képelemeket. Ez a fajta megközelítés különösen célravezetőnek tűnik a több száz, nagyrészt a látható tartományon kívül eső felvételezési sáv adatainak értelmezéséhez.

Távérzékelési alkalmazásokban is elterjedt klaszterezési algoritmus a K-means, ami alapelve szerint a felhasználó által előre meghatározott számú „átlagvektort” jelöl ki a pontthalmazon, melyek a közeli elemek többször ismételt hozzárendelésével és az átlagértékek újraszámolásával gravitálnak a nagyobb pontsűrűségű adattartományok felé (Ball and Hall, 1967). Módosított változata az ISODATA eljárás, mely az osztályokat méretüktől és az

elemek hasonlóságától függően egyesítheti vagy külön klaszterekre is bonthatja az iterációs folyamat során (Schowengerdt 2007).

A belvízzel elöntött, illetve túlnedvesedett talajok térképezésére az irányítatlan osztályozási módszerek tűnnek a legalkalmasabbnak, hiszen az egyes felszíntípusokat gyakran a terepen is nehéz egymástól elkülöníteni, képi klaszterezési algoritmusokkal viszont azonosíthatjuk és csoportosíthatjuk a homogén foltokat.

A képkivágot ISODATA eljárásnak vetettem alá, maximálisan három iterációval és 95%-os konvergencia küszöbvel. A szűkebb mintaterület sajátosságait figyelembe véve a lehetséges klaszterek számát 3 és 10 közé állítottam be. Az adatosztályozás végeredményét a 34. ábra mutatja be. Az algoritmus végül 4 klasztert azonosított, melyek viszonylag egyszerűen megfeleltethetők a főbb belvítérképezési kategóriáknak: a vízzel borított felszíneknek, a nedves és a száraz talajoknak, illetve a növényzetnek.



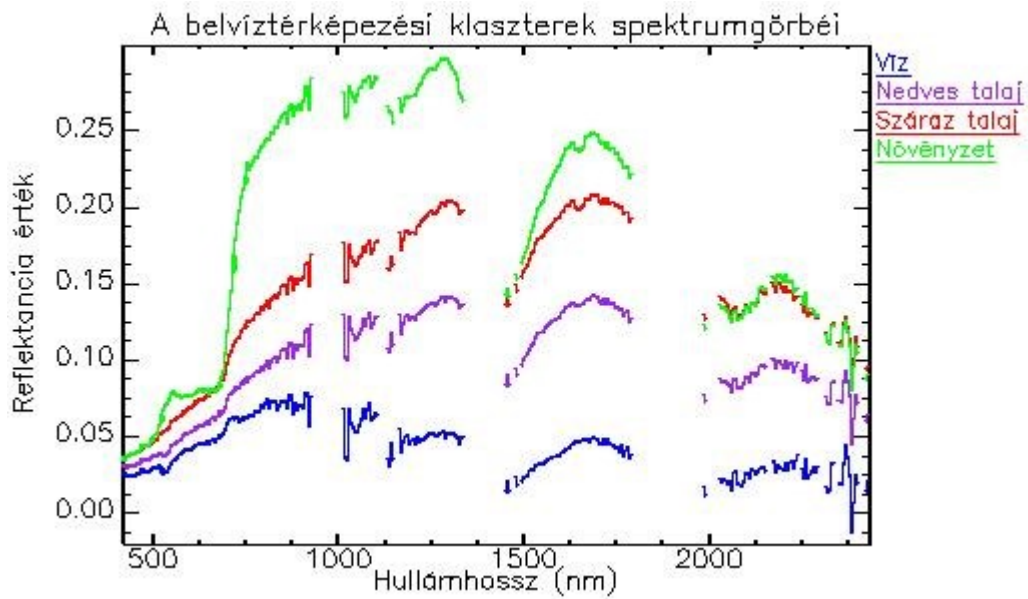
34. ábra: Az ISODATA osztályozás eredménye: vízzel borított (kék), túlnedvesedett (lila), száraz (piros) talajfelszínek és a növényzet (zöld) osztályai

Szeeparabilitási táblázatban foglaltam össze a klaszterezéssel kapott négy osztály képelemeinek spektrális átfedéseit (15. táblázat). Míg a vegetációval fedett felszínek reflektancia értékeit aránylag könnyű elválasztani a többi térképezési kategóriától, a Transformed Divergence-szel számított jelzőszámok 1.8-nal magasabbak, addig a belvízelöntéssel különböző mértékben érintett talajfelszínek, a száraz, a túlnedvesedett és a vízzel teljesen elborított területek osztályozási határvonalai elmosódottabbak, megbízhatóan csak a két végletet jelentő klaszterek, a vízmentes talajok és a nyílt vízfoltok különíthetők el (Jeffries-Matusita: 1.924, Transformed Divergence: 1.98). Figyelemre méltó különbség mutatkozik a két index eredményei között a száraz talajfelszínek és a növényzet viszonylatában, a Jeffries-Matusita módszer számottevően érzékenyebb (1.271) az osztályok között fennálló, vélhetően a felszántott területen visszamaradt növényi maradványok okozta spektrális keveredésből fakadó átfedésekre, mint a Transformed Divergence (1.879).

Jeffries-Matusita	Vízfelszín	Túlnedvesedett talaj	Száraz talaj	Növényzet	
Növényzet	1,98	1,828	1,271	0	Növényzet
Száraz talaj	1,924	1,222	0	1,879	Száraz talaj
Túlnedvesedett talaj	1,477	0	1,273	1,999	Túlnedvesedett talaj
Vízfelszín	0	1,716	1,98	2	Vízfelszín
	Vízfelszín	Túlnedvesedett talaj	Száraz talaj	Növényzet	Transformed Divergence

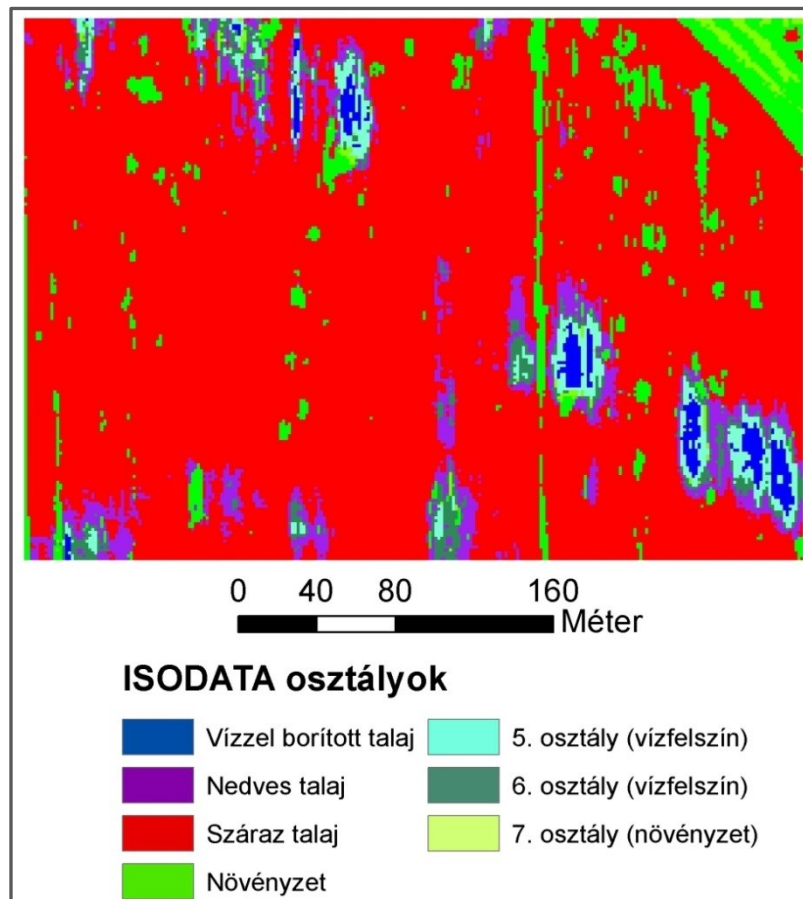
15. táblázat: Az ISODATA eljárással előállított osztályok szeeparabilitása

A klaszterek zajos sávoktól megtisztított reflektancia-görbéit vizsgálva megállapítható, hogy elkülönítésük alapját elsősorban intenzitásbeli különbségeik jelenthették (35. ábra). A nyílt vízhez, illetve a növényzethez köthető osztályok átlagspektrumain jól látszódnak a két felszíntípus jellegzetességei, a vegetációs vörös él, valamint az árnyékos és a vizes területek alacsony függvényértékeinek abszcisszáét követő lefutása. A kétféle talajtípus közti eltérést azonban szinte kizárólag a képelemek fényessége, azaz a görbék vertikális pozíciója képezi.



35. ábra: Az ISODATA eljárással előállított osztályok spektrumgörbéi

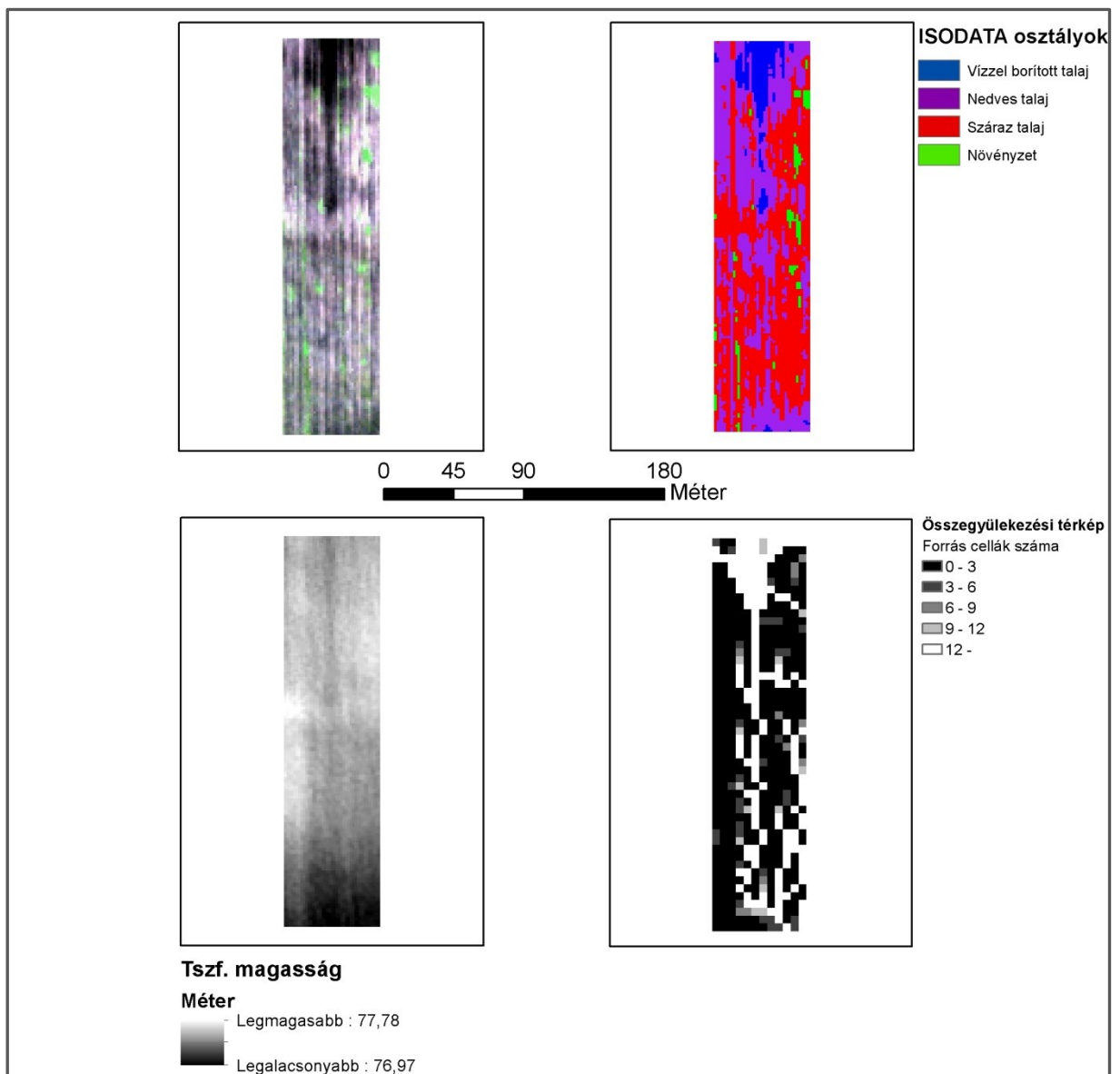
Az előző fejezetben az irányítatlan osztályozást nem normalizált adaton végeztem el. Bár a reflektancia értékek normalizációjával kiküszöbölhetjük a megvilágítás eltéréseiből adódó képi anomáliákat, a belvíztérképezésnél mégsem feltétlenül célszerű ezt a korrekciót alkalmazni. Terepi felmérések híján a nedves (telített) és a száraz talajokat valószínűleg vagy színes infravörös felvételek alapján szokás elkülöníteni (van Leeuwen 2012), a két kategória között pedig gyakran csak árnyalatnyi eltérések vannak, ezért a normalizáció könnyen el is simíthatja a különbségeket. A 36. ábra mutatja a korábbiakkal megegyező beállításokkal, de ezúttal normalizált adaton lefuttatott ISODATA osztályozás eredményét. A korábbi 4 osztály mellett 3 újabb, vízfelszíneket és növényzetet magában foglaló klasztert azonosított az eljárás, amit világoskék, tengerzöld és sárgászöld (chartreuse) színnel jelöltem. A vízzel telített talajok jelentős része azonban a száraz felszínekkel egy osztályba került, ami főleg a terület északnyugati sarkában szembetűnő (36. ábra)



36. ábra: Az ISODATA osztályozás eredménye normalizált adaton

4.4.7 Osztályozási eredmények összevetése a LiDAR domborzatmodellel

A Természeti Földrajzi és Geoinformatikai Tanszék LiDAR (Light Detection and Ranging) felvételei 2009. november 19-én készültek Optech ALTM 3100 típusú szkennelével, 1500 méteres magasságból a Tápai-rét és Algyő külterülete fölött, mintegy 70 négyzetkilométernyi kiterjedésben, pontsűrűségük 1,4 darab/négyzetméter (Szatmári et al. 2012). Keleti szegélye mentén kb. 60-100 méter szélességben átnyúlik a belvízelöntés mintaterülete fölé, itt lehetőségem nyílt összevetni a lézerszkennelt adatot a hiperspektrális vizsgálatok eredményével (37. ábra).



37. ábra: Az összehasonlító kivágat a hiperspektrális felvételen, valószínűs megjelenítésben (bal felső ábra), a LiDAR domborzatmodellen (bal alsó ábra), az ISODATA osztályozás eredménytérképén (jobb felső ábra), a lefolyási irányokból számított összegyülekezési térkép (jobb alsó ábra)

A lehetséges belvízképződési helyek lehatárolására a mintaterület lejtési viszonyait vizsgáltam meg a nagy pontosságú domborzatmodellen. Hogy a lokális egyenetlenségeket kiszűrjem és az eredményt könnyebben értelmezhetővé tegyem, a raszteres adat felbontását egyről 5 méteresre módosítottam, majd a lefolyási irányokat a már módosított geometriájú modellen számítottam ki. Összegyülekezési térképet (flow accumulation modell) az egyes

képelemekhez rendelt folyásirányi szögek alapján lehet összeállítani, értékei pedig a vízfolyás forrásaiként megjelölhető, magasabban elhelyezkedő pixelek számát mutatják.

Észak-dél irányú barázda rajzolódik ki a domborzati térképen (37. ábra), melyet követnek az összegyülekezési modell magasabb, fehérrel jelzett értékei is. A hosszanti mélyedésen kívül a kivágot északi és déli szélén is találunk olyan felszíneket, ahol a lefolyási becslések alapján jelentősebb akkumuláció várható, ezek a területek ugyanakkor az ISODATA módszerrel készített felszínborítási térképen is megjelennek, a nedves, illetve a vízzel borított talajok kategóriájába sorolva.

	Teljes mintaterület (7770 pixel)	Nyílt víz (752 pixel)	Nedves talaj (3121 pixel)	Száraz talaj (3567 pixel)	Növényzet (330 pixel)
Minimum	76,88	76,88	76,91	76,93	76,94
Maximum	77,78	77,57	77,71	77,78	77,75
Átlag	77,412	77,268	77,396	77,451	77,469
Szórás	0,208	0,248	0,219	0,171	0,203

16. táblázat: Az ISODATA belvíztérképezési osztályok domborzati adatai (t. sz. f. m., méter)

A hiperspektrális és a LiDAR domborzati adatok összehasonlítására kijelölt, szűkebb kivágoton azonosítottam az ISODATA klaszterezés osztályait, majd kiszámítottam az egyes felszíntípusok főbb magassági paramétereit (Csendes & Mucsi 2016). Statisztikai eredményeim alapján a terület magasabban fekvő részeit a száraz talajok foglalják el viszonylag alacsony szórás mellett (16. táblázat). A nedvesebb, illetve nyílt vízborítású felszínek értékei ezzel szemben, bár átlagaik (77,396 és 77,268 m) a teljes kivágoténál (77,412 m) alacsonyabbak, nagyobb változékonyságot mutatnak. Következtetésként megállapítható, hogy a belvízelöntés térbeli mintázatát kis lépték és néhány tíz cm-es magasságkülönbségek esetén is elsősorban a mikrorelief, a kisebb lokális egyenetlenségek határozzák meg, ennek oka pedig a vizet kevésbé áteresztő, agyagos talajszerkezetre vezethető vissza.

4.5 A statisztikai módszerek alkalmazása városi mintaterületeken

4.5.1 Statisztikai vizsgálatok városi mintaterületen

Az SZTE Természeti Földrajzi és Geoinformatikai Tanszékén a városi felszínek térképezését és a magassági információkat tartalmazó felszínmodellek készítését célzó vizsgálatokba a hiperspektrális adatokat is bevontuk. Feladatom az volt, hogy a mintaterületen, Szeged délnyugati városrészeiben felvett tanítóterületeket a helyszínen ellenőrizsem, majd a spektrumkönyvtárakba rendezett mintapontokon a képosztályozásokat előkészítendő spektrális szeparabilitási számításokat végezzek. További feladatom volt, hogy a spektrális értékeiben nagy változékonyságot mutató, jelentős zajtartalmú városi légi felvételek tanítóterületeinek jellemzésére az osztályok homogenitását számszerűsítő statisztikai módszerek találjak.

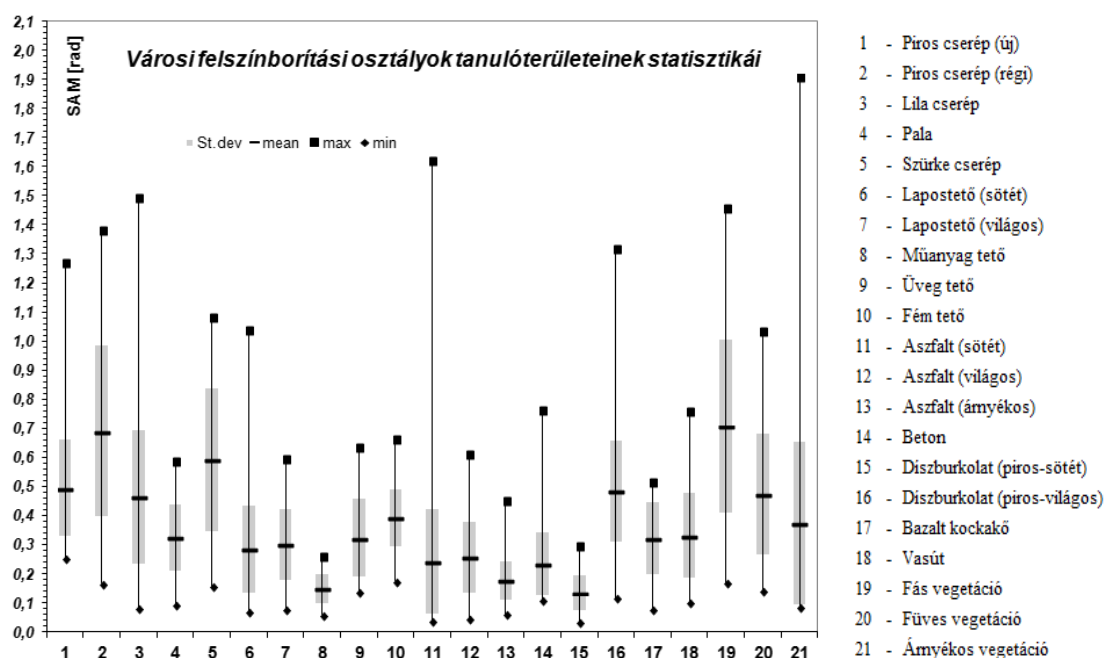
Az átfedések vizsgálatára Jeffries-Matusita-féle szeparabilitási számítást és szögeltérés-alapú spektrális indexet alkalmaztam, mely 0 és 1 közötti értékkel, relatív skálán jelzi a spektrális hasonlóságot, ahol 1 a teljes átfedésnek felel meg. Eredményeink alapján alacsonyabb spektrális távolság az egyes tetőtípusok spektrumgörbéi között található, továbbá az aszfalt és a beton felszínek mutatnak egymás felé nagyobb hasonlóságot (17. táblázat), ami a képosztályozások százalékos pontossági értékeiben is visszaköszönt (Tobak et al. 2012a).

A városi tanítóterületek spektrális mintáira is alkalmaztuk az általam kifejlesztett szögeltérési box plot diagramot, melyen a spektrális homogenitást az osztályhoz tartozó képelemek átlagspektrummal bezárt spektrális szöge jellemzi (4.2.7 fejezet). Mintaterületünk felszíntípusai közül az égetett tetőcserepek és a növényzet spektrális heterogenitása magasabb (38. ábra), az egyéb tetőburkolatok (pala, műanyag, fém, üveg) értékei ellenben jóval szűkebb tartományban szóródnak, ami pontosabb osztályozási eredményeikre is magyarázattal szolgál (Tobak et al. 2013).

	Régi piros tető-cserép	Lila tető-cserép	Új szürke tető-cserép	Új piros tető-cserép	Lakótelepi beton felszín	Fák, cserjék	Fű	Aszfalt	Bazalt díszkő	JM index
Régi piros	-	1,454	1,971	1,677	1,964	1,99	1,999	1,99	1,99	Régi piros
Lila	0,708	-	1,764	1,592	1,963	1,999	1,999	1,946	1,99	Lila
Új szürke	0,701	0,564	-	1,985	1,973	1,999	2	1,712	1,925	Új szürke
Új piros	0,935	0,715	0,681	-	1,999	1,999	2	1,997	1,999	Új piros
Lakótelep	0,677	0,53	0,89	0,655	-	2	2	1,8	1,822	Lakótelep
Fa, cserje	0,142	0	0,013	0,122	0,018	-	1,99	1,999	1,999	Fa, cserje
Fű	0,17	0	0,046	0,149	0,052	0,925	-	2	1,999	Fű
Aszfalt	0,64	0,397	0,756	0,624	0,786	0,181	0,22	-	1,282	Aszfalt
Bazalt	0,519	0,258	0,623	0,5	0,653	0,264	0,304	0,845	-	Bazalt
SAM index	Régi piros tető-cserép	Lila tető-cserép	Új szürke tető-cserép	Új piros tető-cserép	Lakótelepi beton felszín	Fák, cserjék	Fű	Aszfalt	Bazalt díszkő	

17. Táblázat: A városi felszínborítási osztályok spektrális elkülöníthetősége

(Tobak et al. 2012a)



38. ábra: A városi felszínborítási osztályok box plot statisztikája spektrális szögeltérés felhasználásával (Tobak et al. 2013)

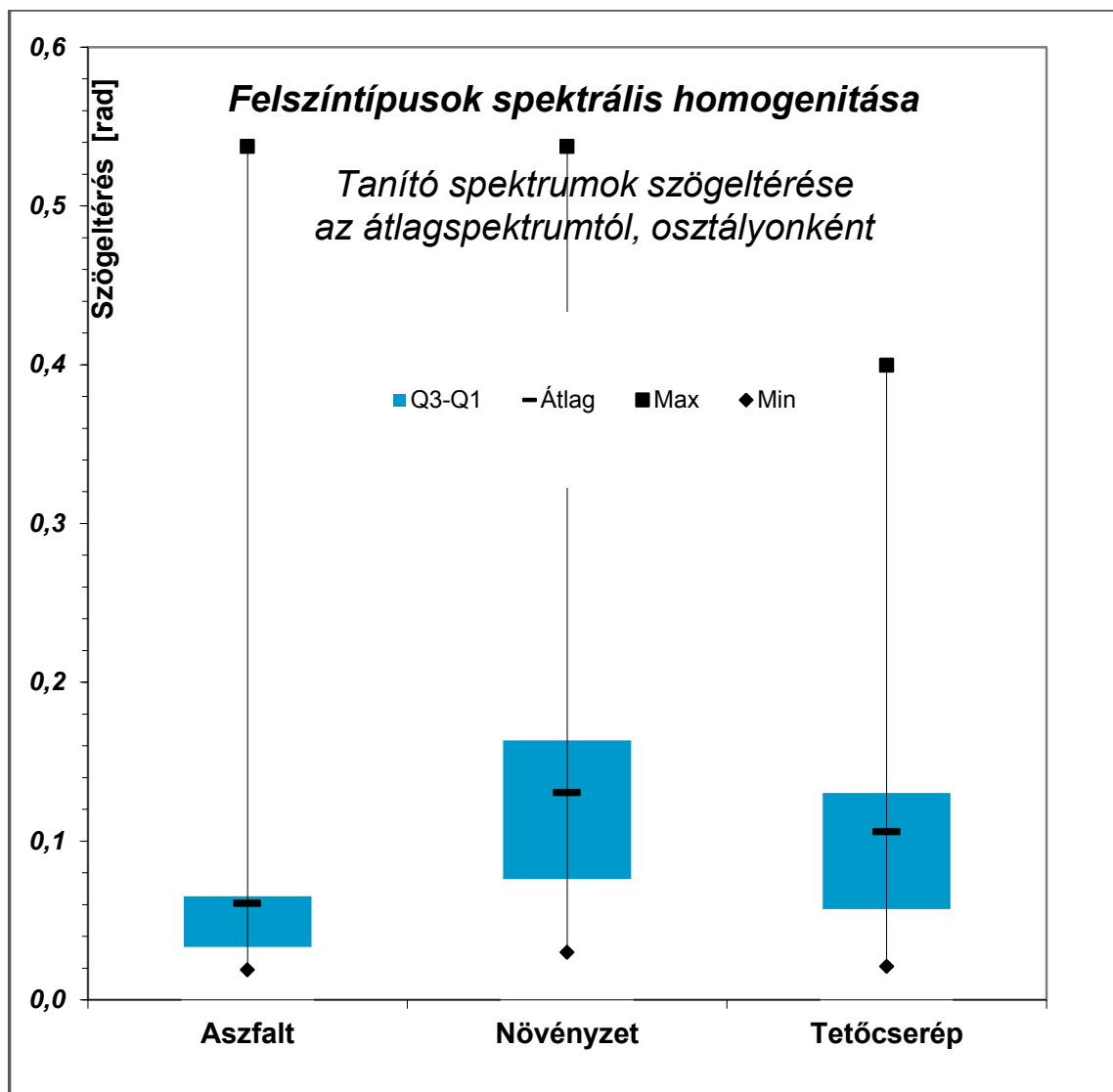
4.5.2 Kertvárosi térképezés laboratóriumi spektrumok alapján

A 2010 szeptemberében készített hiperspektrális felvételek Szeged-Alsóváros belterületét lefedő kivágatáról tematikus felszínborítási térképet készítettünk szerzőtársaimmal (Mucsi et al. 2016). A laboratóriumi spektrumok bemérése az SZTE Orvosi Fizikai és Orvosi Informatikai Intézetében, AVANTES spektrométerrel, AVANTES DH2000-es fényforrás segítségével történt. Három, jellegzetes kertvárosi felszín típusról (tetőcserép, aszfalt, növényzet) gyűjtöttünk spektrális mintákat a légi felvételeken, melyek közül az osztályozáshoz használt tanítóterületeket a laboratóriumi spektrumok segítségével, spektrális szögeltérés alapján válogattuk le. A tanítópontok közé azok a képelemek kerültek, melyek reflektancia értékei alacsonyabb szögeltérést mutattak a spektrométerrel bemért minták átlagértékétől, ezek további statisztikai vizsgálata az én feladatomból volt.

	Aszfalt	Növényzet	Tetőcserép	TD
Aszfalt	0	2	1,88	Aszfalt
Növényzet	1,97	0	1,7	Növényzet
Tetőcserép	1,63	1,48	0	Tetőcserép
JM	Aszfalt	Növényzet	Tetőcserép	

18. táblázat: A kertvárosi tanítóterületek Jeffries-Matusita (JM) és Transformed Divergence (TD) szeparabilitási értékei

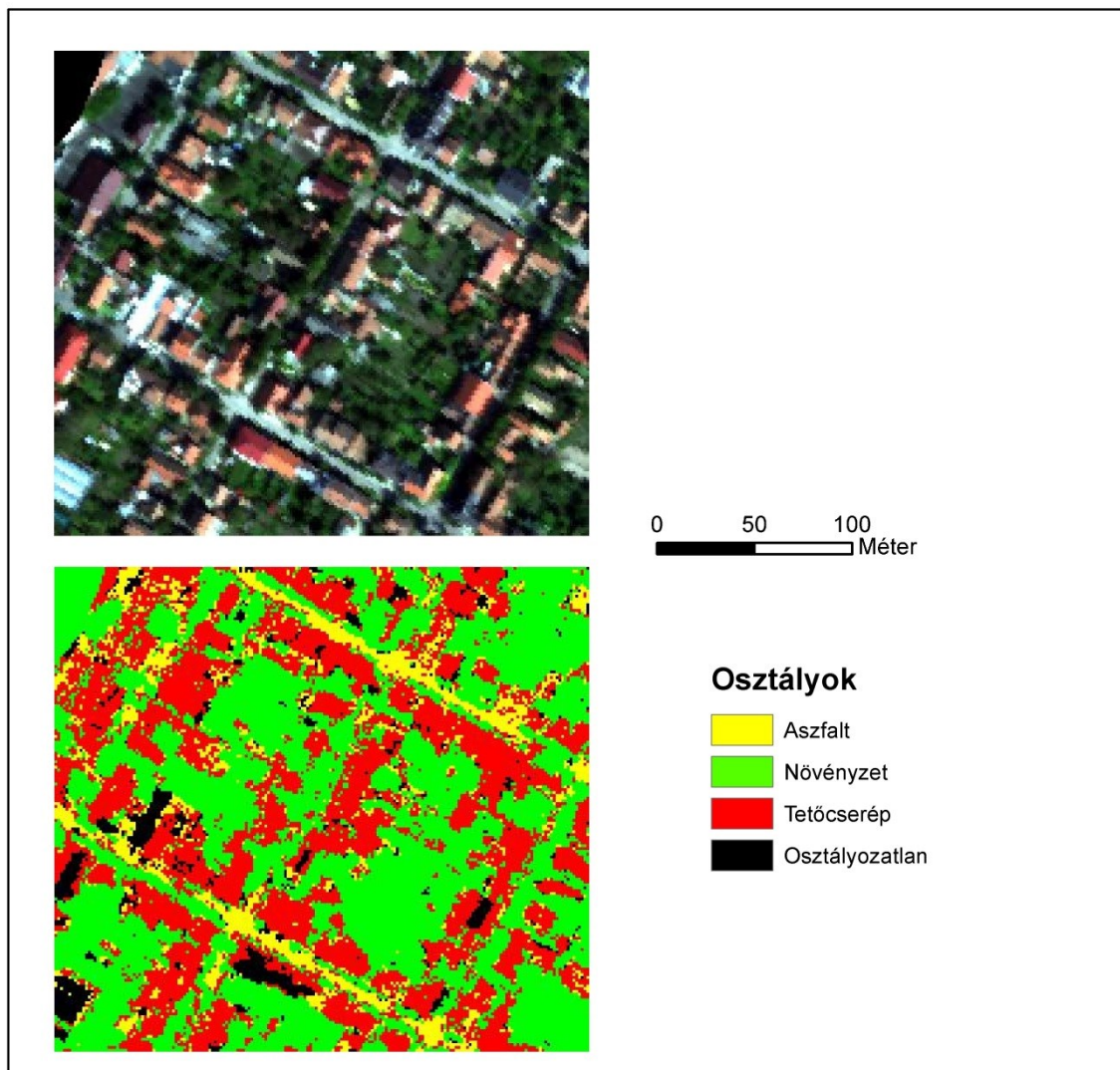
Az osztályok elkülöníthetőségét 0-tól 2-ig terjedő szeparabilitási indexszámokkal jellemeztem (18. táblázat). Alacsonyabb, átfedésre utaló értékek a tetőcserép és a növényzet (Jeffries-Matusita 1,48, Transformed Divergence 1,7), valamint a tetőcserép és az aszfalt (JM 1,63, TD 1,88) viszonylatában jelentkeznek, ami alapján feltételezhetjük, hogy a tetőcserepek osztályának tanítóterülete egyes képelemeket is tartalmaz. A vegetáció és az aszfalt felszín indexszámai a távérzékelés irodalmában alkalmazott 1,9-es határérték felett maradnak, így köztük nem kellett számottevő félreosztályozásokkal számolni.



40. ábra: A kertvárosi osztályok spektrális homogenitása szögeltérésük alapján (Box plot diagramon)

A három vizsgált felszínborítási osztály spektrális homogenitását az átlagos szögeltérésük alapján hasonlítottam össze (40. ábra). Az aszfalt osztály képi értékeit néhány kiugró, maximálisan mintegy 0,55 radián eltérést mutató elemet leszámítva az átlagértéket és a negyedek közti különbséget tekintve nagyfokú homogenitás jellemzi. A vegetáció pixeleinek reflektanciája ezzel szemben, mivel a tanítóterületek több növényfaj elegyeit is magukban foglalják, jóval nagyobb szórású, mindazonáltal jellegzetes spektrumgörbéjének köszönhetően a növénytakaró pontos osztályozása nem jelent különösebb nehézséget. Magas átlagos spektrális szögeltérése (0,1 rad) és alacsony szeparabilitási értékei (18. táblázat) folytán a tetőcserepek osztálya már kitettebb az átfedéseknek, félreosztályozásoknak.

Szerzőtársaimmal hatvanhét hiperspektrális sáv alapján térképeztük a mintaterületet, melyeket, összhangban a laboratóriumi mérésekkel, az 500 és 800 nanométer közötti tartományból vettünk. Osztályozásunkat Maximum likelihood módszerrel végeztük, a három kategória egyikébe se tartozó képelemek kiszűrésére pedig 90%-os probabilitási küszöbértéket határoztunk meg. Az osztályozatlan pixelek mindenekelőtt a nem cserépből készült, illetve az erősebben megvilágított, a képeken zajosan megjelenő tetőknél találhatók (41. ábra). Az eredménytérképen sikerült elkülöníteni a három vizsgált felszíntípust, a teljes pontosság valamennyi keresztvalidációs pontthalmazon 95% fölötti, az aszfalt burkolat azonosítását ugyanakkor néhány esetben akadályozta az utcák fölé belógó lombkorona.



41. ábra: A kertvárosi mintaterület valószínűségi kivágata (felső kép), illetve a Maximum likelihood osztályozás eredménye (alsó kép). Forrás: Mucsi et al. (2016)

5. Összefoglalás

A földrajztudomány kutatási témái közé éppúgy hozzátartozik a környezeti változások térbeliségének, földrajzi mintázatának vizsgálata, mint a felszíni formák kialakulásának leírása. Az adatgyűjtési technológia fejlődésével, a térbeli vonatkozású adatok bővülésével nem minden esetben tud lépést tartani a beérkezett információtartalom feldolgozása, annak kiértékelése során korábban kevésbé ismert földrajzi sajátosságok kerülhetnek előtérbe, ami újfajta megközelítések és tudományos paradigma kialakítására ösztönzi a kutatókat. Különösen látványos a képi és egyéb távérzékelt adatok szélesebb körű alkalmazása és minőségi javulása, ami részben a magasabb fokú automatizációval függ össze.

Jelen dolgozatom tárgya a nagy spektrális felbontású, több száz sávban felvételező képalkotó spektrométerek, más néven hiperspektrális szenzorok adatsorainak felhasználása földrajzi-földtudományi vizsgálatokban. Mintaterületemen, a Szeged mellett található Tápai-réten viszonylag kis területen több, magyarországi tájakra jellemző környezeti probléma is megfigyelhető, köztük a belvízképződés, az özönnövények terjedése, illetve a szénhidrogén-bányászattal járó környezeti károk is, hiszen itt működik a MOL egyik termelési egysége. Nagyfelbontású adataim egyedülálló lehetőséget nyújtottak a felszíni anyagok vizsgálatára és térképezésére, a megbízható spektrális diszkriminációhoz azonban több esetben módszertani fejlesztésekre, újszerű statisztikai megközelítésekre volt szükség.

A hiperspektrális képek AISA szenzorral készültek, másfél méteres geometriai és 4,29 – 6,28 nanométeres spektrális felbontással, összesen 359 sávban. A felvételezésre 2010 szeptember 22-én került sor, a vegetációs időszak végén. A megelőző hónapok különösen csapadékosak voltak, a vizsgált kistáj és a Dél-Alföld jelentős részén kiterjedt belvízelöntések jelentkeztek (Szatmári & van Leeuwen 2013), ami alkalmat adott a túlnedvesedett talajfelszínek távérzékeléses vizsgálatára is. Terepi felméréseimet 2010 tavasza és 2013 között végeztem, ennek során részletes katasztert készítettem az ártéri élőhelyek növényfajairól, földhivatali források alapján begyűjtöttem a helyi gazdálkodóktól a mezőgazdasági művelés 2010-es adatait, valamint azonosítottam a mintaterületen található összes belvízfoltot.

Zajszűrést végeztem a légi felvételeken, melynek során a hibás sávokat a képelemek területi autokorrelációs értékei alapján válogattam ki. A további vizsgálatokhoz

főkomponens-transzformációt végeztem, illetve a reflektancia értékek intenzitásbeli különbségeit adatnormalizációval mérsékeltem.

A kijelölt tanítóosztályok spektrális homogenitásának számszerűsítésére egy új módszert fejlesztettem ki: a mintapontok szórását az átlaguktól számított spektrális szögeltérés szerint értelmeztem. Az így nyert statisztikai értékeket box plot diagramon mutattam be, ezen kívül véletlenszerűen kiválasztott spektrális csatornák osztályokra vetített szórásával jellemeztem a vizsgált felszíntípusokat.

Tanítóterületeim spektrális szeeparabilitását Jeffries-Matusita és Transformed Divergence eljárással számítottam ki a főkomponens sávokon, 0-tól 2-ig terjedő indexértékekkel jellemeztem a minták közti átfedéseket, amivel a várható félreosztályozásokon túl a térképezési algoritmusok esetleges túlilleszkedését is előre tudtam jelezni. Az osztályok elkülöníthetőségét a szeeparabilitási indexek mellett a statisztikai T-próbával is megmértem, melynek során a gyalogakác és a szürke nyár reflektancia görbéit sávonként hasonlítottam össze, továbbá scatter plot diagramon ábrázoltam a főbb felszíntípusok elhelyezkedését a spektrális térben.

Összehasonlítottam a távérzékelés nemzetközi szakirodalmában elterjedt képosztályozási eljárások eredménytérképeit pontosságuk és megbízhatóságuk szempontjából. Az elfogultságtól mentes kiértékelés érdekében az eredeti tanító-ellenőrző pontkiosztáson túl további 5 keresztvalidációs halmazt generáltam Matlab programmal, melyek átlagos pontosságát és szórását is figyelembe vettem. A szögeltérési térképezés (SAM) átlagosan a vizsgált képelemek csupán 64,5 százalékát osztályozta helyesen, a nem-paraméteres, spektrális átlagszámításon alapuló módszer különösen a természetes, nem mezőgazdasági növényfajok esetén mutatott magas hibaarányt. A paraméteres eljárások közül a Maximum likelihoodal hajtottam végre osztályozást, amihez az input sávok számát főkomponens transzformációval csökkentettem. Eredménytérképe magas, valamennyi keresztvalidációs halmazon 92% fölötti teljes pontosságot nyújtott, megjelenített képrészletén (20. ábra) azonban néhány osztály, elsősorban az amerikai köris és a piros tetőcserepek dominanciája figyelhető meg, ami a gépi tanulási algoritmusok túlilleszkedésére utalhat a kevés mintán. Overfitting probléma szerényebb mértékben az SVM osztályozásnál is jelentkezett, a vizsgált módszerek közül ugyanakkor itt volt legalacsonyabb a hibaarány (kappa-index: 0,97, teljes pontosság: 97%).

A Rice Egyetem statisztikai laboratóriumában hibrid SOM (Self-Organizing Map) vizsgálatot végeztem, az irányítatlan gépi tanulási folyamatához pedig a normalizált felvételeket használtam fel. Az így előállt topológiai mátrix klasztereit (23. ábra) irányított módon címkéztem föl, a művelet magában foglalta a súlyértékek korrekcióját is. Pontossága némileg elmarad a többi gépi tanulási módszer eredményeitől (86,03% az eredeti pontthalmazon), ami részben a klaszterezésbe bevont képkivágat felszín típusainak nem teljes körű ismeretével és az így kihagyott osztályok címkéinek hiányával magyarázható. Irányítatlan tanulásának köszönhetően az SOM eljárás felismerte az egymáshoz hasonló képelemeket és osztályozatlanul hagyta a tanítópontok klasztereihez nem rendelhető pixeleket, növelve ezáltal az eredménytérkép megbízhatóságát is.

Tanítóterületeken végzett statisztikai számításokkal előre tudtam jelezni az osztályozási hibákat, az első öt főkomponens alapján számított, 1,8 alatti Jeffries-Matusita index jelezte a szignifikáns spektrális átfedéseket, a kevésbé homogén, referenciaterületein túlterjeszkedni hajlamos térképezési kategóriákat, mint az amerikai köris (A) vagy a piros tetőcserepek (F), pedig a 0,08 radián fölötti átlagos szögeltérésű osztályok között találtam (16. ábra). A pontosság növelésén és a mintapontok megválasztásán túl a képi statisztikák önmagukban is alkalmasak a felszín típusok spektrális karakterisztikájának, ezáltal anyagi minőségének jellemzésére is.

Irányított és irányítatlan osztályozásokkal térképeztem a belvízzel érintett mezőgazdasági parcellákat, előbbieik közül a Spectral Angle Mapping 84, a Binary Encoding pedig 90 százalékos pontossággal különítette el a túlnedvesedett és a száraz talajokat, míg az irányítatlan ISODATA klaszterezéssel sikerült valamennyi ellenőrzőpontot elkülöníteni. Az eredménytérképek alapján ugyanakkor megállapítottam, hogy a tanítóterületek alkalmazásán alapuló megközelítés a belvíztérképezésben nem feltétlenül célravezető, a túlnedvesedett felszínnek összetett térbeli mintázata és képelemeinek nagyfokú spektrális változatossága megnehezíti azonosításukat, ezért az előzetes terepi adatgyűjtést nem igénylő irányítatlan osztályozások bizonyulnak megbízhatóbbnak a hiperspektrális képek kiértékelésére. Szeparabilitási indexekkel hasonlítottam össze az ISODATA eljárással meghatározott négy klaszter spektrális elkülöníthetőséget, jelentős átfedést jelző alacsony értéket a száraz és a túlnedvesedett talajokat reprezentáló osztályok között fedeztem fel (1,271-es Jeffries-Matusita index), ami szintén a két felszín típus szétválasztásának bizonytalanságával hozható összefüggésbe. A domborzati viszonyokat LiDAR modell segítségével jellemeztem, a

belvízelöntés előrejelzésére a relatív magasságkülönbségek alapján számított összegyülekezési térkép (Flow Accumulation Map) bizonyult alkalmasnak.

A spektrális homogenitást és szeparabilitást városi felszínekre is kiszámítottam a képosztályozások előkészítéséhez, továbbá a laboratóriumi mérések alapján kiválogatott tanítóterületek spektrális elkülöníthetőségét támasztottam alá statisztikai méréseimmel.

Kitekintő fejezetemben megállapítottam, hogy nagyobb térbeli léptékű osztályozásokra, környezetföldrajzi monitoring-vizsgálatokra is alkalmasak a hiperspektrális szenzorral felszerelt, nagy időfelbontással felvételező műholdak, mint az USGS EO-1 vagy a közeljövőben fellőni tervezett német EnMAP képalkotó rendszere. A közeljövőben várhatóan előtérbe kerülnek azok a távirányítású, pilóta nélküli távérzékelési eszközök is, köztük a drónok és a quadrocopterek, melyektől a légi vizsgálatok költségének zuhanását és a képi adatbázisok ugrásszerű bővülését remélhetjük.

Irodalomjegyzék

- Adam, E. & Mutanga, O. (2009): Spectral discrimination of papyrus vegetation (*Cyperus papyrus* L.) in swamp wetlands using field spectrometry. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64, pp. 612-620.
- Al-Moustafa, T. Armitage, R.P., & Danson, F.M. (2012): Mapping fuel moisture content in upland vegetation using airborne hyperspectral imagery. *Remote Sensing of Environment*, 127, pp. 74-83.
- Andrew, M.E. & Ustin, S. L. (2008): The role of environmental context in mapping invasive plants with hyperspectral image data. *Remote Sensing of Environment*, 112, pp. 4301-4317.
- Anne, N.J.P., Abd-Elrahman, A.H., Lewis, D.B., & Hewitt, N.A. (2014): Modeling soil parameters using hyperspectral image reflectance in subtropical coastal wetlands. *International Journal of Applied Earth Observation and Geoinformation*, 33, pp. 47-56.
- Bachmann, Ch.M., Philpot, W., Abelev, A., & Korwan, D. (2014): Phase angle dependence of sand density observable in hyperspectral reflectance. *Remote Sensing of Environment*, 150, pp. 53-65.
- Ball, G. H., & Hall, D. J. (1967): A clustering technique for summarising multivariate data. *Behavioral Science*, 12, pp. 153-155.
- Bartholy, J., Pongrácz, R., & Gelybó, Gy. (2007): Regional climate change expected in Hungary for 2071-2100. *Applied ecology and environmental research*, 5 (1), pp. 1-17.
- Belluco, E., Camuffo, M., Ferrari, S., Modenese, L., Silvestri, S., Marani, A., & Marani, M. (2006): Mapping salt-marsh vegetation by multispectral and hyperspectral remote sensing. *Remote Sensing of Environment*, 105, pp. 54-67.
- Benediktsson, J. A., Sveinsson, J. R., & Arnason, K. (1994): Classification of very-high-dimensional data with geological applications. In: *Proc. MAC Europe*, 91, pp. 13-18
- Borengasser, M., Hungate, W. S., & Watkins, R. (2008): *Hyperspectral Remote Sensing: Principles and Applications*. Taylor & Francis Group.
- Bota, V. (2004): Inváziós növények elleni tevékenységek a nemzeti park igazgatóságokban – Körös-Maros Nemzeti Park. In: Mihály, B. & Botta-Dukát, Z. (eds.): *Özönnövények – Biológiai inváziók Magyarországon*, TermészetBÚVÁR Alapítvány Kiadó, Budapest, pp. 405-406.
- Botta-Dukát, Z. (2004): A növényi invázióval kapcsolatos hazai és nemzetközi aktivitás. In: Mihály, B. & Botta-Dukát, Z. (eds.): *Özönnövények – Biológiai inváziók Magyarországon*, TermészetBÚVÁR Alapítvány Kiadó, Budapest, pp. 17-33.
- Burai, P. (2006): Földhasználat-elemzés és növény-monitoring különböző adattartalmú és térbeli felbontású távérzékelési felvételek alapján. *Agrártudományi Közlemények 2006/22 Különszám*, pp. 7-12.
- Burai, P. (2007): Távérzékelés IV. Oktatási segédanyag, HEFOP 3.3.1

- Burai, P. & Tomor, T. (2011): Hiperspektrális felmérés eredményei az Ipoly Balassagyarmat és Ipolytarnóc közötti szakaszán. A Bükk Nemzeti Park Igazgatóság természeti értékeinek kutatása I.: „Az Ipoly-vízgyűjtő vizes élőhelyeinek komplex felmérése, közösségi jegyzékeinek kidolgozása” (Záró konferencia és záró értekezés), *Felsőtárkány, 2011. február 10-11.*
- Burai, P., Lövei, G. Zs., Lénárt, Cs., Nagy, I., & Enyedi, P. (2010): Mapping aquatic vegetation of the Rakamaz-Tiszanagyfalui Nagy-morotva using hyperspectral imagery. *Acta Geographica Debrecina Landscape & Environment*, 4 (1), pp. 1-10.
- Burai, P., Deák, B., Oroly, V., & Lénárt, Cs. (2014): Mapping of Grass Species Using Airborne Hyperspectral Data. *International Workshop on Remote Sensing and GIS for Monitoring of Habitat Quality (Vienna, 24-25 September 2014, Proceedings)*, pp. 87-88.
- Büttner, Gy. (1990): Digitális légifelvételezési kampány a Szigetközben, Földfelszíni és meteorológiai megfigyelések a világűrben, *MANT*, pp. 105-112.
- Büttner, Gy., Csillag, F., & Mather, P. M. (1988): Spectral And Spatial Information Content Of Spot Data. In: *Geoscience and Remote Sensing Symposium, 1988. IGARSS '88. Remote Sensing: Moving Toward the 21st Century.*, International; 10/1988
- Cawley, G. C. & Talbot, N. L. C. (2010): On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation. *Journal of Machine Learning Research*, 11, pp. 2079-2107.
- Campbell, J.B. (1981): Spatial correlation effects upon accuracy of supervised classification of land cover. *Photogrammetric Engineering and Remote Sensing*, 47, pp. 355-357.
- Chauhan, P., Kaur, P., Srivastava, N., Sinha, R. K., Jain, N., & Murty, S. V. S. (2015): Hyperspectral remote sensing of planetary surfaces: an insight into composition of inner planets and small bodies in the solar system. In: *Current Science, vol. 108, No. 5 Special section: Hyperspectral Remote Sensing*, pp. 915-924.
- Chiu, H. Y. & Collins, W. (1978): A spectroradiometer for airborne remote sensing. *Photogramm. Eng. And Remote Sensing* 44, pp. 507-517.
- Chiuderi, A. and Cappellini, V., 1996, A Kohonen's self-organising map for land cover classification. In: Parlow, E. (ed.): *Progress in Environmental Remote Sensing Research and Applications*, Proceedings of the 15th EARSeL Symposium, Basle, Switzerland, 4-6 September 1996, A.A. Balkema, Rotterdam, pp. 107-112.
- Clark, M. L., Roberts, D. A., & Clark, D. B. (2005): Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sensing of Environment*, 96, pp. 375-398.
- Clevers, J. P. G. W. (1994): Imaging spectrometry in agriculture - plant vitality and yield indicators. In: Hill, J. & Mégier, J. (eds.): *Imaging Spectrometry: a Tool for Environmental Observations*. Kluwer Academic Publishers.
- Cohen, J. (1960): A coefficient of agreement form nominal scales, *Education Psychological Measurement*, Vol. 20, No. 1, pp. 37-46.
- Comelli, D., Nevin, A., Gulotta, D., Toniolo, L., Valentini, G., & Cubeddu, R. (2008): Integrated hyperspectral and time resolved fluorescence imaging combined with statistical

data analysis: diagnostic investigations of wall paintings. 9th International Conference on NDT of Art, Jerusalem Israel, 25-30 May 2008

Congalton, R. C. & Green, K. (2009): *Assessing the Accuracy of Remotely Sensed Data. Principles and Practices*. Taylor & Francis Group.

Cortes, C. & Vapnik, V. N. (1995): Support vector networks. *Machine Learning*, vol. 20, no. 3, pp. 273-297, Sept. 1995.

Cortijo, F. J. & Perez de la Blanca, N. (1999): The performance of regularised discriminant analysis versus non-parametric classifiers applied to high dimensional image classification. *International Journal of Remote Sensing*, 20, pp. 3345-3365.

Csendes, B. (2010): *Belvíztérképezés szubpixel-alapú osztályozással, Landsat 7 ETM+ műholdfelvételek alapján*. Diákköri dolgozat, XXX. Országos Tudományos Diákköri Konferencia, Nyíregyháza

Csendes, B. (2012): Detection of invasive plants on the flood plain of river Tisza, using hyperspectral airborne imagery. In: Malvic, T., Geiger, J., & Cvetkovic, M. (eds.): *Geomathematics as geoscience*, 4th Croatian-Hungarian and 15th Hungarian geomathematical congress, pp. 187-194.

Csendes, B. (2013): Invazív növények spektrális tulajdonságainak vizsgálata légifotókon. In: Keresztes, G. (ed.): *Tavaszi Szél*, 2013: Spring wind, pp. 408-414.

Csendes, B. & Mucsi, L. (2016): Inland excess water mapping using hyperspectral imagery. *Geographica Pannonica*, Vol. 20, Issue 4, *in print*.

Csiszár, Á & Bartha, D. (2004): Amerikai köris. In: Mihály, B. & Botta-Dukát, Z. (eds.): *Özönnövények – Biológiai inváziók Magyarországon*, TermészetBÚVÁR Alapítvány Kiadó, Budapest, pp. 131-142.

Deák, M., Ónodi, Zs., & Mari, L. (2012): Magyarországi falvak lehatárolása EO-1 - hiperspektrális űrfelvételek segítségével. In: Lóki, J. (ed.): *Az elmélet és gyakorlat találkozása a térinformatikában III. - Térinformatikai konferencia és szakkiallítás*. Konferencia helye, ideje: Debrecen, Magyarország, 2012.05.24-2012.05.25. Debrecen: Debreceni Egyetemi Kiadó, 2012. pp. 119-126.

Deák, M., Telbisz, T., Árvai, M., Mari, L., & Horváth, F. (2014): Subpixel Vegetation Classification of EO-1 Hyperion Data Through Spectral Reduction, A Case Study in Hungary. In: Yingbin, H. Mingjie, G., & Zhenya, Z. (eds.): *The Application of Remote Sensing and GIS Technology in Crop Production*. Konferencia helye, ideje: Beijing, 2013.08.26-2013.08.30. China: China Agricultural Science and Technology Press (CASTP), 2014. pp. 83-88.

Dövényi, Z. (ed.) (2010): *Magyarország kistájainak katasztere*. MTA Földrajztudományi Kutatóintézet.

Fairweather, S., Potter, Ch., Crabtree, R., & Li, S. (2012) A Comparison of Multispectral ASTER and Hyperspectral AVIRIS Multiple Endmember Spectral Mixture Analysis for Sagebrush and Herbaceous Cover in Yellowstone. *Photogrammetric Engineering & Remote Sensing*, 78, 23-33.

- Fernandes, M. R., Aguiar, F. C., Silva, J. M. N., Ferreira, M. T., & Pereira, J. M. C. (2013): Spectral discrimination of giant reed (*Arundo donax* L.): A seasonal study in riparian areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 80, pp. 80-90.
- Geary, R. C. (1954): The Contiguity Ratio and Statistical Mapping. *The Incorporated Statistician*, 5 (3), pp. 115-145.
- Gleeson, D. F., Pappalardo, R. T., Grasby, S. E., Anderson, M. S., Beauchamp, B., Castano, R., Chien, S. A., Doggett, T., Mandrake, L., & Wagstaff, K. L. (2010): Characterization of a sulfur-rich Arctic spring site and field analog to Europa using hyperspectral data. In: *Remote Sensing of Environment*, 114, pp. 1297-1311.
- Goetz, A. F. H., Vane, G., Solomon, J. E., & Rock, B. N. (1985): Imaging Spectrometry for Earth Remote Sensing *Science*, vol. 228, no. 4704, pp. 1147-1153.
- Goetz, A. F. H. (1991): Imaging Spectrometry for studying Earth, Air, Fire and Water, *EARSel, Advances in Remote Sensing*, 1, 3-15.
- Green, R., Eastwood, M., Sarture, C., Chrien, T., Aronsson, M., Chippendale, B., Faust, J., Pavri, B., Chovit, C., Solis, M., & Olah, M. (1998): Imaging Spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). *Remote Sensing of Environment*, 65 (3), p. 227.
- Gualtieri, J. A. & Cromp, R. F. (1999): Support Vector Machines for Hyperspectral Remote Sensing Classification. *Proc. SPIE*, 3584, pp. 221-232.
- Haggett, P. (2001): *Geography: A Global Synthesis*. Pearson Education.
- Hargitai, H. (2006): *A hiperspektrális képfeldolgozás módszerei és az első magyarországi képkalkotó spektrométeres repülés adatainak elemzése*. PhD értekezés, ELTE.
- Hestir, E.L., Khanna, S., Andrew, M.E., Santos, M.J., Viers, J.H., Greenberg, J.A., Rajapakse, S.S., & Ustin, S.L. (2008): Identification of invasive vegetation using hyperspectral remote sensing in the California Delta ecosystem. *Remote Sensing of Environment*, 112, pp. 4034-4047.
- Horler, D.N.H., Dockray, M., & Barber, J. (1983): The red edge of plant leaf reflectance. *International Journal of Remote Sensing*, 4, pp. 273-288.
- Howell, E. S., Merényi, E., & Lebofsky, L. A. (1994): Classification of asteroid spectra using a neural network. *J. Geophys. Res.* 99 (E5), pp. 10847-10865.
- Huang, C., Davis, L. S. & Townshend, J. R. G. (2002): An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, 23, pp. 725-749.
- Hughes, G. F. (1968): On the mean accuracy of statistical pattern recognizers, *IEEE Transactions on Information Theory*, Vol. IT-14, pp. 56-63.
- Hunter, P.D., Tyler, A.N., Carvalho, L., Codd, G.A., & Maberly S.C. (2010): Hyperspectral remote sensing of cyanobacterial pigments as indicators for cell populations and toxins in eutrophic lakes. *Remote Sensing of Environment*, 114, pp. 2705-2718.

- Inoue, Y., Sakaiya, E., Zhu, Y., & Takahashi, W. (2012): Diagnostic mapping of canopy nitrogen content in rice based on hyperspectral measurements. *Remote Sensing of Environment*, 126, pp. 210-221.
- Jay, S. & Guillaume, M. (2014): A novel maximum likelihood based method for mapping depth and water quality from hyperspectral remote-sensing data. *Remote Sensing of Environment*, 147, pp. 121-132.
- Jensen, J. R. (1986): *Introductory digital image processing: A remote sensing perspective*. Englewood Cliffs, NJ: Prentice-Hall.
- Jia, X. & Richards, J. A. (1993): "Binary coding of imaging spectrometry data for fast spectral matching and classification." *Remote Sensing of Environment*, 43, pp. 47-53.
- Johnston, R. (1997): *Introduction to microbiotic crusts*. United States Department of Agriculture Natural Resources Conservation Service, http://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/nrcs142p2_053263.pdf
- Joshi, C., de Leeuw, J. , Skidmore , A. K. , van Andel, J., Lekhak, H.D., & van Duren, I. C. (2005): Remote sensing and GIS for mapping and management of invasive shrub *C. odorata* in Nepal. *AGILE 2005, 8th conference on Geographic Information Science, 26-28, May, 2005, Portugal*, pp. 71-80.
- Jung, A., Kardeván, P., & Tókei, L. (2005): Detection of urban effect on vegetation in a less build-up hungarian city by hyperspectral remote sensing, *Physics and Chemistry of the Earth*, 30, pp. 255-259.
- Kardeván, P., Vekerdy, Z., Róth, L., Sommer St.–Kemper, Th., Jordan, GY., Tamás, J., Pechmann, I., Kovács, E., Hargitai, H., & László, F. (2003): Outline of scientific aims and data processing status of the first Hungarian hyperspectral data acquisition flight campaign, *3rd EARSeL Workshop on Imaging Spectroscopy in Oberpfaffenhofen*, 2003. május 13-16.
- Kertész, P., Király, G., & Burai, P. (2014): Tree Species Mapping Using Airborne Hyperspectral Remote Sensing. *International Workshop on Remote Sensing and GIS for Monitoring of Habitat Quality (Vienna, 24-25 September 2014, Proceedings)*, pp. 60-62.
- Knox, N.M., Skidmore, A.K., Prins, H.H.T., Asner, G.P., van der Werff, H.M.A., de Boer, W.F., van der Waal, C., de Knegt, H.J., Kohi, E.M., Slotow, R., & Grant, R.C. (2011): Dry season mapping of savanna forage quality, using hyperspectral Carnegie Airborne Observatory sensor. *Remote Sensing of Environment* 115, pp. 1478-1488.
- Kohonen, T. (1997): *Self-Organizing Maps, 2nd ed.* Springer, Heidelberg
- Kokaly, R.F., Couvillion, B.R., Holloway, J.M., Roberts, D.A., Ustin, S.L., Peterson S.H., Khanna, S., & Piazza, S.C. (2013): Spectroscopic remote sensing of the distribution and persistence of oil from the Deepwater Horizon spill in Barataria Bay marshes. *Remote sensing of Environment*, 129, pp. 210-230.
- Kozma-Bognár, V. & Berke, J. (2010): New Evaluation Techniques of Hyperspectral Data. *Journal of Systemics, Cybernetics and Informatics*, 5 (8), pp. 49-53.
- Kozma-Bognár, V. & Berke, J. (2013): Többsávós légifelvételek optimális adatsávjainak meghatározása. In: Czúni, L. (ed.): *A Képfeldolgozók és Alakfelismerők Társaságának 9.*

országos konferenciája - KÉPAF 2013. Konferencia helye, ideje: Bakonybél, Magyarország, 2013.01.29-2013.02.01. Veszprém: NJSZT-KÉPAF, pp. 105-119.

Kruse, F. A., Lefkoff, A. B., Boardman, J. W., Heidebrecht, K. B., Shapiro, A. T., Barloon, J. P., & Goetz, A. F. H. (1993): The spectral image processing system (SIPS) - Interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, 44, pp. 145-163.

Labovitz, M.L. & Matsuoko, E.J. (1984): The influence of autocorrelation on signature extraction – an example from a geobotanical investigation of Cotter Basin, Montana. *International Journal of Remote Sensing*, 5, pp. 315-332.

Landgrebe, D. A. (2003): *Signal Theory Methods in Multispectral Remote Sensing*. John Wiley & Sons, Inc.

Lary, D. J., Alavi, A. H., Gandomi, A. H., & Walker, A. L. (2015): Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, 7, pp. 3-10.

Lausch, A., Salbach, C., Doktor, D., Schmidt, A., Merbach, I., & Pause, M. (2015): Deriving phenology of barley with imaging hyperspectral remote sensing. *Ecological Modelling* 295, pp. 123-135.

Lausch, A., Bannehr, L., Beckmann, M., Boehm, C., Feilhauer, H., Hacker, J. M., Heurich, M., Jung, A., Klenke, R., Neumann, C., Pause, M., Rocchini, D., Schaepman, M. E., Schmidtlein, S., Schulz, K., Selsam, P., Settele, J., Skidmore, A. K., & Cord, A. F. (2016): Linking Earth Observation and taxonomic, structural and functional biodiversity: Local to ecosystem perspectives. *Ecological indicators*. doi: 10.1016/j.ecolind.2016.06.22. (in press)

Lee, K. S., Cohen, W. B., Kenned, R. E., Maersperger, T. K., & Gower, S. T. (2004): Hyperspectral versus multispectral data for estimating leaf area index in four different biomes. *Remote sensing of Environment*, 91, pp. 508-520.

van Leeuwen, B. (2012): *Artificial neural networks and geographic information systems for inland excess water classification*. PhD értekezés, SZTE.

van Leeuwen, B., Mezösi, G., Tobak, Z., Szatmári, J., & Barta, K. (2012): Identification of inland excess water floodings using an Artificial Neural Network. *Carpathian Journal of Earth and Environmental Sciences*, November 2012, Vol. 7, No. 4, pp. 173 -180.

van Leeuwen, B., Henits, L., Mészáros, M., Szatmári, J., Tobak, Z., Pavic, D., Savic, S., & Dolinaj, D. (2013a): Belvíz-elöntések lehatárolása RapidEye műhold-felvételek alapján. *Hidrológiai Közlöny*, 93 (3), pp. 17-24.

van Leeuwen, B., Henits, L., Mészáros, M., Tobak, Z., Szatmári, J., Pavic, D., Savic, S., & Dolinaj, D. (2013b): Classification methods for inland excess water modeling. *Journal of Environmental Geography*, 6 (1-2), pp. 1-8.

ifj. Lele, J. (1978): A Tápairét tanyái. In: *A Móra Ferenc Múzeum évkönyve 1976/77-I*, pp. 253-281.

Lewis, S.A., Robichaud, P.R., Frazier, B.E., Wu, J.Q., & Laes D.Y.M. (2008): Using hyperspectral imagery to predict post-wildfire soil water repellency. *Science Direct Geomorphology*, 95, pp. 192-205.

- Licciardi, G., Marpu, P. R., Chanussot, J., & Benediktsson, J. A. (2012): Linear versus nonlinear PCA for the classification of hyperspectral data based on the extended morphological profiles. *IEEE Trans. Geosci. Rem. Sens.* 9 (3), pp. 447-451.
- Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (2004): *Remote Sensing and Image Interpretation (Fifth Edition)*. John Wiley & Sons, Inc.
- Liu, L., Wang, B., & Zhang L. (2010a): An approach based on self-organizing map and fuzzy membership for decomposition of mixed pixels in hyperspectral imagery. *Pattern Recognition Letters*, 31, pp. 1388-1395.
- Liu, Zh-Y., Wu, H-F., & Huang, J-F. (2010b): Application of neural networks to discriminate fungal infectio levels in rice panicles using hyperspectral reflectance and principal components analysis. *Computers and Electronics in Agriculture* 72, pp. 99-106.
- Markham, B. M. & Barker, J. L. (1986): Landsat MSS and TM Post-Calibration Dynamic Ranges, Exoatmospheric Reflectances and At-Satellite Temperatures. EOSAT Landsat Technical Notes, no. 1, Earth Observation Satellite Co., Lanham, MD, USA.
- Mather, P. M. (2004): *Computer Processing of Remotely-Sensed Images: An Introduction*. John Wiley & Sons Ltd.
- Matheson, D.S. & Dennison, Ph. E. (2012): Evaluating the effects of spatial resolution on hyperspectral fire detection and temperature retrieval. *Remote sensing of Environment*, 124, pp. 780-792.
- Mazer, A. S., Martin, M., Lee, M., & Solomon, J. E. (1988): Image Processing Software for Imaging Spectrometry Analysis. *Remote Sensing of the Environment*, v. 24, no. 1, pp. 201-210.
- van der Meer, F. D. (2004): Pixel-Based, Stratified and Contextual Analysis of Hyperspectral Imagery. In: de Jong, S. M. & van der Meer, F. D. (eds.): *Remote Sensing Image Analysis: Including the Spatial Domain*, Dordrecht, Springer Academic Publishers, pp. 153-180.
- van der Meer, F. D., de Jong, S. M., & Bakker, W. (2001): Imaging spectrometry: Basic analytical techniques. In: van der Meer, F. D. & de Jong, S. M. (eds.): *Imaging spectrometry – Basic Principles and Prospective Applications*, Kluwer Academic Publishers, Dordrecht, pp. 17-64.
- Merényi, E., Farrand, W.H., & Tracadas, P. (2004): Mapping Surface Materials on Mars From Mars Pathfinder Spectral Images With HYPEREYE. *Proc. International Conference on Information Technology (ITCC 2004), April 5-7, 2004, Las Vegas, NV, USA. vol II*, pp 607-614.
- Merényi, E., Csathó, B., & Tasdemir, K. (2007a): Knowledge discovery in urban environments from fused multi-dimensional imagery. *Proc. 4th IEEE GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas (URBAN 2007)*, Paris, France, April 11-13, 2007. (invited paper). pp 1-13. DOI: 10.1109/URS.2007.371860 , IEEE Catalog number 07EX1577.
- Merényi, E., Farrand, W. H., Brown, R. H., Villmann, T., & Fyfe, C. (2007b): Information extraction and knowledge discovery from high-dimensional and high-volume complex data sets through precision manifold learning, *Proc. NASA Science Technology Conference (NSTC2007)*, College Park, Maryland, June 19 - 21, 2007, 11pp. ISBN 0-9785223-2-X

- Merényi, E., Tasdemir, K., & Zhang, L. (2009): Learning Highly Structured Manifolds: Harnessing the Power of SOMs. *Similarity-Based Clustering – Lecture Notes in Computer Science*, Volume 5400, pp. 138-168.
- Merényi, E., Farrand, W. H., Taranik, J. V., & Minor, T. B. (2014): Classification of hyperspectral imagery with neural networks: comparison to conventional tools. *EURASIP Journal on Advances in Signal Processing*, 2014, (71) doi:10.1186/1687-6180-2014-71.
- Metternicht, G. I. & Zinck, J. A. (1998): Evaluating the information content of JERS-1 SAR and Landsat TM data for discrimination of soil erosion features. *ISPRS Journal of Photogrammetry & Remote Sensing*, 53, pp. 143-153.
- Montello, D. & Sutton, P. (2013): *An Introduction to Scientific Research Methods in Geography and Environmental Studies*. SAGE Publications Ltd, p. 62.
- Moon, T. & Merényi, E. (1995): Classification of hyperspectral images using wavelet transforms and neural networks. In *Proc. Annual SPIE Conf.* (San Diego, 1995)
- Moshou, D., Vrindts, E., De Ketelaere, B., De Baerdemaeker, J., & Ramon, H. (2001): A neural network based plant classifier. *Computers and Electronics in Agriculture*, 31, pp. 5-16.
- Mucsi, L. (2001): Characterisation of oil-industrial contamination using aerial and thermal images. In: Buchroithner, M. F. (ed.): *Proceedings of the 20th EARSel symposium Dresden: A Decade of Trans-European Remote Sensing Cooperation*, Balkema, Rotterdam, 373-377.
- Mucsi, L. & Henits, L. (2011): Belvívelöntési térképek készítése közepes felbontású ürfelvételek szubpixel alapú osztályozásával. *Földrajzi közlemények*, 135 (4.), pp. 365-378.
- Mucsi, L., Liska, Cs., Tobak, Z., Csendes, B., Nagy, L., & Henits, L. (2016): Mapping built up areas with hyperspectral images and laboratory measurements. *Publikáció alatt*.
- Mutanga, O. & Kumar, L. (2007): Estimating and mapping grass phosphorus concentration in an African savanna using hyperspectral image data. *International Journal of Remote Sensing*, 28, pp. 4897-4911.
- Mutanga, O. & Skidmore, A. K. (2004): Integrating imaging spectroscopy and neural networks to map grass quality in the Kruger National Park, South Africa. *Remote Sensing of Environment*, 90, pp. 104-115.
- O'Neill, A.L. (1994): Reflectance spectra of microphytic soil crusts in semi-arid Australia. *International Journal of Remote Sensing*, 15 (3), pp. 675-681.
- Pal, M & Foody, G. M. (2010): Feature selection for classification of hyperspectral data by SVM. *IEEE Transactions on Geoscience and Remote Sensing*, 48 (5). pp. 2297-2307.
- Pal, M., and P. M. Mather. 2005. Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, 26, pp. 1007-1011.
- Pálfai, I. (2001): A belvível definíciói. *Vízügyi közlemények*, 83. évf. 3. füzet, pp. 376-392.
- Pause, M., Schweitzer, C., Rosenthal, M., Keuck, V., Bumberger, J., Dietrich, P., Heurich, M., Jung, A. & Lausch, A. (2016): In Situ/Remote Sensing Integration to Assess Forest Health- A Review. *Remote Sensing* 8 (471)

- Peng, J., Shen, H., & Wu, J.S. (2013): Soil moisture retrieving using hyperspectral data with the application of wavelet analysis. *Environ. Earth Sci.*, 69 (1), pp. 279-288
- Penn, B. S. (2005): Using self-organizing maps to visualize high-dimensional data. *Computers & Geosciences*, 31, pp. 531–544.
- Platt, U. & Stutz, J. (2008): *Differential Optical Absorption Spectroscopy: Principles and Applications*. Springer Science & Business Media.
- Potter, Ch., Li, S., Huang, S., & Crabtree, R.L. (2012): Analysis of sapling density regeneration in Yellowstone National Park with hyperspectral remote sensing data. *Remote Sensing of Environment* 121, pp. 61-68.
- Priszter Sz. (1997): A magyar adventívflóra kutatása. *Botanikai Közlemények*, 84, pp. 25-32.
- Pysek, P., Prach, K., Rejmánek, M., & Wade, M. (eds.) (1995): *Plant invasions - General aspects and special problems*. SPB Academic Publ., Amsterdam.
- Rakonczai J. (2013): A belvív. In: Szabó, J.: *Általános természetföldrajz I*. Budapest, pp. 221-226. ISBN 978 963 312 062 0
- Rakonczai, J., Farsang, A., Mezősi, G., & Gál, N. (2011): A belvízképződés elméleti háttere. *Földrajzi Közlemények* 135 (4), pp. 339-350.
- Reddy, M. A. (2008): *Textbook of Remote Sensing and Geographical Information Systems (Third Edition)*. BS Publications, Hyderabad.
- Rejmánek, M. (1996): A theory of seed plant invasiveness: the first sketch. *Biological Conservation* 78, pp. 171-181.
- Richards, J.A. (1986): *Remote Sensing-Digital Analysis*. Springer-Verlag, Berlin.
- Richards, J. A. & Jia, X. (2006): *Remote Sensing Digital Image Analysis (Fourth Edition)*. Springer-Verlag Berlin Heidelberg.
- Robert, P. C. (1997): Remote Sensing: A Potentially Powerful Technique for Precision Agriculture. In: *Proceedings: Land Satellite Information in the Next Decade II: Sources and Applications*, American Society for Photogrammetry and Remote Sensing, Bethesda MD, pp 19-25.
- Santini, F., Alberotanza, L., Cavalli, R.M., & Pignatti, S. (2010): A two-step optimization procedure for assessing water constituent concentrations by hyperspectral remote sensing techniques: An application to the highly turbid Venice lagoon waters. *Remote Sensing of Environment* 114, pp. 887-898.
- Santos, M.J., Greenberg, J.A., & Ustin, S.L. (2010): Using hyperspectral remote sensing to detect and quantify southeastern pine senescence effects in red-cockaded woodpecker (*Picoides borealis*) habitat. *Remote Sensing of Environment* 114, pp. 1242-1250
- Schaale, M., & Furrer, R. (1995): Land surface classification by neural networks. *International Journal of Remote Sensing*, 16, pp. 3003-3031.
- Schmidt, K. S. & Skidmore, A. K. (2003): Spectral discrimination of vegetation types in a coastal wetland. *Remote Sensing of Environment*, 85, pp. 92-108.

- Schowengerdt, R. A. (2007): *Remote Sensing: Models and Methods for Image Processing (Third Edition)*. Elsevier Inc.
- Serpico, S. B., Moser, G., & Cattoni, A. F. (2007): Feature reduction for classification purpose. In: Chang, C. (ed.): *Hyperspectral Data Exploitation – Theory and Applications*, John Wiley & Sons Publication, Hoboken, New Jersey, pp. 245-274.
- Shaw G. & Manolakis, D. (2002): Signal Processing for hyperspectral image exploitation, *IEEE Signal Processing Magazine*, vol. 19, no. 1, p. 12.
- Simis, S. G. H. & Olsson, J. (2013): Unattended processing of shipborne hyperspectral reflectance measurements. *Remote Sensing of Environment*, 135, pp. 202-212.
- Skidmore, A. (2002): *Environmental Modelling with GIS and Remote Sensing*. Taylor & Francis, pp. 1-5.
- Skidmore, A. K., Turner, B. J., Brinkhof, W., & Knowles, E. (1997): Performance of a neural network: Mapping forests using GIS and remotely sensed data. *Photogrammetric Engineering And Remote Sensing*, 63, pp. 501-514.
- Sobrino, J. A., Franch, B., Mattar, C., Jiménez-Muñoz, J. C., & Corbari, C. (2012): A method to estimate soil moisture from Airborne Hyperspectral Scanner (AHS) and ASTER data: Application to SEN2FLEX and SEN3EXP campaigns. *Remote Sensing of Environment*, 117, pp. 415-428.
- Somers, B., Cools, K., Delalieux, S., Stuckens, J., Van der Zande, D., Verstraeten, W.W., & Coppin P. (2009): Nonlinear Hyperspectral Mixture Analysis for tree estimates in orchards. *Remote Sensing of Environment*, 113, pp. 1183-1193.
- Somers, B. & Asner, G. P. (2013): Multi-temporal hyperspectral mixture analysis and feature selection for invasive species mapping in rainforests. *Remote Sensing of Environment*, 136, pp. 14-27.
- Somlyódy, L., Nováky, B., & Simonffy, Z. (2010): Éghajlatváltozás, szélsőségek és Vízgazdálkodás, In: Csete, L. (ed.): „Klíma-21” füzetek klímaváltozás – hatások – válaszok 61, MTA KSZI Klímavédelmi kutatások koordinációs iroda, Budapest.
- St. Clair, L.L. & Johansen, J.R. (1993): Introduction to the symposia on soil crust communities. *Great Basin Naturalist*, 53(1), pp. 1-4.
- Stamenkovic, J., Tuia, D., De Morsier, F., Borgeaud, M., & Thiran, J.P. (2013): Estimation of Soil Moisture from Airborne Hyperspectral Imagery with Support Vector Regression. In *Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)* (No. EPFL-CONF-185220).
- Story, M. & Congalton, R. (1986): Accuracy assessment: A user's perspective. *Photogrammetric Engineering and Remote Sensing*. Vol. 52, No. 3, pp. 397-399.
- Stuart, J. (2012): *Spectral Unmixing Applied to Desert Soils for the Detection of Sub-Pixel Disturbances*. Naval Postgraduate School, Monterey.
- Szatmári, J., Szíjj, N., Mucsi, L., Tobak, Z., van Leeuwen, B., Lévai, Cs., & Dolleschall, J. (2012): Comparing LIDAR DTM with DEM-5 of Hungary. In: Geiger, J., Pál-Molnár, E., & Malvić, T. (eds.): *New horizons in Central European geomathematics, geostatistics and*

geoinformatics: *Selected studies of the 2011 Croatian-Hungarian Geomathematical Convent*, Mórahalom. pp. 151-158.

Szatmári, J. & van Leeuwen, B. (eds.) (2013): *Inland excess water – Belvíz – Suvišne unutrašnje vode*. Szeged – Novi Sad, ISBN 978-963-306-263-0

Szigetvári, Cs. & Tóth, T. (2004): Gyalogakác. In: Mihály, B. & Botta-Dukát, Z. (eds.): *Özönnövények – Biológiai inváziók Magyarországon*, TermészetBÚVÁR Alapítvány Kiadó, Budapest, pp. 187-206.

Thenkabail, P.S., Smith, R.B., & De Pauw, E. (2000): Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment*, 71, pp. 158-182

Thomas, V., Treitz, P., Jelinski, D., Miller, J., Lafleur, P., & McCaughey, J. H. (2002): Image classification of a northern peatland complex using spectral and plant community data. *Remote Sensing of Environment*, 84, 83-99.

Tits, L., Somers, B., Stuckens, J., Farifteh, J., & Coppin, P. (2013): Integration of in situ measured soil status and remotely sensed hyperspectral data to improve plant production system monitoring: Concept, perspectives and limitations. *Remote sensing of Environment*, 128, pp. 197-211

Tobak, Z., Csendes, B., Henits, L., van Leeuwen, B., & Mucsi, L. (2012a): A városi felszín spektrális tulajdonságainak vizsgálata légifelvételek alapján. In: Nyári, D. (ed.): *Kockázat - Konfliktus - Kihívás: A VI. Magyar Földrajzi Konferencia, a MERIEXWA nyitókonferencia és a Geográfus Doktoranduszok Országos Konferenciájának Tanulmánykötete*, Szeged: SZTE TTK Természeti Földrajzi és Geoinformatikai Tanszék, pp. 1088-1097. ISBN:978-963-306-175-6

Tobak, Z., Csendes, B., Henits, L., van Leeuwen, B., Szatmári, J., & Mucsi, L. (2012b): Városi felszínek spektrális tulajdonságainak vizsgálata légifelvételek alapján. In: Lóki, J. (ed.): *Az elmélet és gyakorlat találkozása a térinformatikában III.*, Debrecen: Debreceni Egyetemi Kiadó, pp. 413-420. ISBN:978-963-318-218-5

Tobak, Z. (2013): *A városi felszín vizsgálata nagy térbeli és spektrális felbontású légifelvételek felhasználásával*. PhD értekezés, SZTE.

Tobak, Z., Csendes, B., Henits, L., van Leeuwen, B., & Mucsi, L. (2013): Légifelvételek spektrális és térbeli információtartalmának felhasználása a városi felszínborítás térképezésében. In: Lóki, J. (ed.): *Az elmélet és a gyakorlat találkozása a térinformatikában IV.*: pp. 441-449. ISBN 978-963-318-334-2

Torrecilla, E., Stramski, D., Reynolds, R. A., Millán-Núñez, E., & Piera, J. (2011): Cluster analysis of hyperspectral optical data for discriminating phytoplankton pigment assemblages in the open ocean. *Remote Sensing of Environment*, 115, pp. 2578-2593.

Tso, B. & Mather, P. M. (2009): *Classification Methods for Remotely Sensed Data (Second Edition)*. Taylor & Francis Group.

Ustin, L. S., Valko, P. G., Kefauver, S. C., Santos, M. J., Zimpfer, J. F., & Smith, S. D. (2009): Remote sensing of biological soil crust under simulated climate change manipulations in the Mojave Desert. *Remote Sensing of Environment*, 113, pp. 317-328.

- Vaiphasa, C., Ongsomwang, S., Vaiphasa, T., & Skidmore, A. K. (2005): Tropical mangrove species discrimination using hyperspectral data: A laboratory study. *Estuarine, Coastal and Shelf Science*, 65, pp. 371-379.
- Vaiphasa, C., Skidmore, A. K., de Boer, W. F., & Vaiphasa, T. (2007): A hyperspectral band selector for plant species discrimination. *ISPRS Journal of Photogrammetry & Remote Sensing*, 62, pp. 225-235.
- Vane, G. (1985): High Spectral Resolution Remote Sensing of the Earth. *Sensors*, 1985, no. 2, pp. 11-20.
- Vane, G., Duval, J. E., & Wellmann, J. B. (1993): Imaging spectroscopy of the Earth and other solar system bodies. In: C. M. Pieters & P. A. J. Englert (eds.): *Remote Geochemical Analyses: Elemental and Mineralogical Composition*, Cambridge University Press, Cambridge, pp. 121-144.
- Vapnik, V. (1998): *Statistical learning theory*. New York, John Wiley.
- Villmann, T., Der, R., Herrmann, M., & Martinetz, T. (1997): Topology Preservation in Self-Organizing Feature Maps: Exact Definition and Measurement. *IEEE Transactions on Neural Networks*, 8(2), pp. 256-266.
- Villmann, T., Merényi, E., & Hammer, B. (2003): Neural maps in remote sensing image analysis. *Neural Networks*, 16, pp. 389-403.
- Visser, F., Wallis, C., & Sinnott, A. M. (2013): Optical remote sensing of submerged aquatic vegetation: Opportunities for shallow clearwater streams. *Limnologia - Ecology and Management of Inland Waters*, 43 (5), pp. 388-398.
- Visi-Rajczi, E., Burai, P., Király, G., & Albert, L. (2012): Ecological Characterization of the green areas in Sopron by Plant Chemical Analysis and Hyperspectral Recording. *International Scientific Conference on Sustainable Development & Ecological Footprint*, March 26-27 2012 Sopron, Hungary
- Warner, T. A. & Shank, M. C. (1997): Spatial Autocorrelation Analysis of Hyperspectral Imagery for Feature Selection. *Remote Sensing of Environment*, 60, pp. 58-70.
- Weber, B., Olechowski, C., Knerr, T., Hill, J., Deutschewitz, K., Wessels, D.C.J., Eitel, B. & Budel, B. (2008): A new approach for mapping of biological soil crusts in semidesert areas with hyperspectral imagery. *Remote Sensing of Environment*, 112, pp. 2187-2201.
- Williams, A.P. & Hunt Jr., E.R. (2002): Estimation of leafy spurge cover from hyperspectral imagery using mixture tuned matched filtering. *Remote Sensing of Environment*, 82, pp. 446-456.
- Yang, C., Everitt, J. H., & Johnson, H. B. (2009): Applying image transformation and classification techniques to airborne hyperspectral imagery for mapping Ashe juniper infestations. *International Journal of Remote Sensing*, 30, pp. 2741-2758.
- Zhang, C. (2015): Applying data fusion techniques for benthic habitat mapping and monitoring in a coral reef ecosystem. *ISPRS Journal of Photogrammetry and Remote Sensing*, 104, pp. 213-223.

Zhang, C. & Xie, Zh. (2012): Combining object-based texture measures with a neural network for vegetation mapping in the Everglades from hyperspectral imagery. *Remote Sensing of Environment*, 124, pp. 310-320.

Zhong, Y. & Zhang, L. (2012): An Adaptive Artificial Immune Network for Supervised Classification of Multi-/Hyperspectral Remote Sensing Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 50, pp. 894-909.

Köszönetnyilvánítás

Ezúton szeretném kifejezni köszönetemet dr. Merényi Erzsébet tanárnőnek a hallatlan lelkesedéssel és munkabefektetéssel végzett témavezetői tevékenységéért. Külön hálával tartozom neki a nemzetközi kutatási pályázatom lebonyolításában való közreműködéséért és segítségéért.

Szeretném továbbá megköszönni dr. Mucsi László témavezetői munkáját, akinek a hiperspektrális távérzékeléses kutatási téma kijelölését és a számtalan fontos kritikai észrevételét is köszönhetem, melyek mind előbbre vitték doktori munkámat.

Köszönettel tartozom houstoni kollégámnak, Patrick O'Driscollnak, aki a képfeldolgozó laboratóriumban segítette munkámat. Tudományos felkészültségével és sokoldalú, gyakorlatias látásmódjával nagyban inspirálta doktori kutatásaimat.

Köszönet illeti dr. Molnár Zsoltot, az MTA Ökológiai Kutatóközpont munkatársát szakmai útmutatásaiért, dr. Deák József Áront és dr. Margóczy Katalint a vegetáció felmérés elméletének és gyakorlati módszereinek megismertetéséért, továbbá László Annát a statisztikai vizsgálatokhoz nyújtott segítségéért.

Szeretném megköszönni dr. Tobak Zalán, dr. Barta Károly, dr. Szilassi Péter, dr. Boudewijn van Leeuwen és dr. Mezősi Gábor szakmai iránymutatásait és jó tanácsait, valamint az SZTE Természeti Földrajzi és Geoinformatikai Tanszékén, az ECFIN főigazgatóságon és a lipcei UFZ intézetben dolgozó volt munkatársaim segítőkészségét és biztatását.

Végül, de nem utolsó sorban szeretném megköszönni családomnak és barátaimnak a sok támogatást, amit dolgozatom elkészítéséhez kaptam.

Summary

Among lots of others, Earth Science research themes include analyzing the spatiality and patterns of environmental changes as well as describing the formation of different surface forms. With data collection methods evolving, the amount of spatial data increases immensely, which does not necessarily result in processing incoming data in due time. Data processing may also reveal less well-known geographical characteristics which encourage researchers to develop new approaches and scientific paradigms. With their ever improving quality it is especially spectacular to use various images and remote sensing data more and more extensively, which is partially due to higher level automatism.

The research theme of my present dissertation is the geographical and environmental application of hyperspectral datasets produced by high spectral resolution spectrometers which produce images in several hundred bands. Tápai Meadow (Tápai-rét) near Szeged is the study area of my research, which is a relatively small area, but it represents a collection of different environmental problems also characteristic of other Hungarian regions, for example excess groundwater, spreading of invasive plant species and environmental damage caused by petroleum extraction, as one of the production units of MOL (Hungarian Oil and Gas Company) operates here. The high resolution data I used provided a unique opportunity to analyze and map the land cover, however, reliable spectral discrimination required methodological developments and new statistical approaches.

Hyperspectral imaging was made by an AISA sensor, including 359 bands altogether, and with 1.5 m geometrical and 4.29-6.28 nm spectral resolutions. Hyperspectral data acquisition took place on 22 September 2010, at the end of the vegetation (growing) period. The previous months were particularly rainy, which resulted in the occurrence of excess groundwater on the surface of a significant part of the Southern Great Plain (Szatmári & van Leeuwen 2013), including the study area. It also made the remote sensing analysis of the oversaturated soil surfaces possible. I collected field data from the spring of 2010 to 2013, which included preparing a detailed register of the plant species of the floodplain ecosystem; based on land registry information, I collected the 2010 data on agricultural production from the local farmers; and I identified all patches of excess groundwater of the study area.

I did noise reduction on the aerial photographs: I separated the imperfect bands on the basis of the spatial autocorrelation values of the pixels during this process. I performed

principal component transformation for further analyses, and I employed data normalization to reduce the differences between the intensity of reflectance values.

I developed a new method to calculate the spectral homogeneity of the designated training areas: I defined the standard deviation of the sample points on the basis of their spectral angle properties. I represented these statistical values on a box plot diagram, in addition, I characterized the studied surface types with the standard deviation of randomly selected spectral channels.

I calculated the spectral separability of my training areas with the Jeffries-Matusita and the Transformed Divergence methods on the principal component bands; 0 to 2 index values were assigned to the overlaps between the patterns, so this way I could predict the possible mis-classifications as well as the possible overfitting of the mapping algorithms. Besides separability indices, I performed a statistical t-test to separate the classes, during which process I compared the reflection curves of desert false indigos (*Amorpha fruticosa*) and grey poplars (*Populus x canescens*) band by band, and I also prepared a scatter plot diagram showing the localization of the main surface types in the spectral space.

I compared the results maps of spectral imaging methods found in the international literature of remote sensing on the basis of their accuracy and reliability. In order to have an unbiased evaluation, I employed MATLAB to prepare another five cross-validation folds besides the original training-testing set of classification points; the average accuracy and standard deviation of these folds were also taken into account. The Spectral Angle Mapper (SAM) classified only an average of 64.5 percent of the studied pixels properly; the non-parametric method based on the calculation of average spectra showed a particularly high error rate in the case of natural, non-agricultural plant species. I chose also the parametric method of Maximum Likelihood for classification, for which I reduced the number of input bands with principal component transformation. Its result map showed an over 92% overall accuracy concerning all cross-validation folds, however, some classes in the depicted pixels (Figure 20) are dominated by green ashes (*Fraxinus pennsylvanica*) and red roof tiles first of all, which may indicate the overfitting of machine learning algorithms. The SVM classification also had a moderate overfitting problem, its error rate was the lowest among the studied methods though (kappa-index: 0.97, overall accuracy: 97%).

I performed hybrid SOM (Self-Organizing Map) analysis in the statistical lab of Rice University, and I used the normalized images for the unsupervised machine learning process. I

applied a supervised method to label the thus created clusters of the topology matrix (Figure 23), and the process also included the correction of weight values. Its accuracy falls behind a little bit compared to the results of other machine-learning methods (86.03% on the original set of points), which is partly due to the fact that the surface types of the pixels involved in clustering were not completely known, therefore some classes were omitted, which also meant a lack of labels. Due to its unsupervised learning, the SOM method recognized the pixels similar to each other, and left those pixels unclassified which could not be assigned to the clusters of the training points, hereby increasing the reliability of the result map.

I could predict classification errors with statistical calculations performed on the training areas, separability values below 1.8 Jeffries-Matusita index, which was calculated on the basis of the first five principal components, indicated significant spectral overlaps. Mapping categories which were less homogeneous and were likely to extend beyond their reference areas such as green ashes (A) or red roof tiles (F) were found among the classes having an average angular deviation of more than 0.08 radians (Figure 16). Besides increasing accuracy and examining sample point selection, the above-mentioned image statistics are also suitable for describing the spectral characteristics of different surface types and their material quality.

I mapped the agricultural parcels affected by excess groundwater by means of supervised and unsupervised classification. Spectral Angle Mapper provided 84% accuracy, while Binary Encoding provided a 90% accuracy (both of them supervised methods) when separating oversaturated and dry soils, while the unsupervised ISODATA clustering separated all test points. However, based on the result maps I ascertained that the method of employing training areas in mapping excess groundwater is not necessarily practical as the complicated spatial patterns of the oversaturated soil surfaces and the great spectral diversity of their pixels make their identification difficult, therefore unsupervised classification methods not requiring previous field data collections prove more reliable for evaluating hyperspectral images. I used separability indices to compare the spectral separability of those four clusters which I had defined with the ISODATA method, and I found low values (1.271 Jeffries-Matusita index) in the classes representing dry and oversaturated soils, which values indicate significant overlaps, that may also be related to the uncertain separability of the two surface types. I characterized topographical relief with the LiDAR model, the Flow Accumulation Map so generated, which I had calculated on the basis of relative height differences, proved to be suitable for predicting excess groundwater.

For preparing image classification, I calculated the spectral homogeneity and separability of urban surfaces too, and my statistical calculations supported the spectral separability of the training areas selected on the basis of laboratory measurements.

In my chapter on future perspectives I ascertained that satellites with hyperspectral sensors and high temporal resolution are capable of providing data for greater scale spatial classification as well as environmental geographical monitoring, for example, USGS EO-1 or the imaging system of the German EnMAP, which is designed for orbit insertion. Remote-controlled, unmanned remote sensing instruments, including drones and quadcopters, are likely to be used more and more frequently in the near future, which may result in the cost reduction of aerial studies and the sudden increase of image databases.

Ábrák és táblázatok jegyzéke

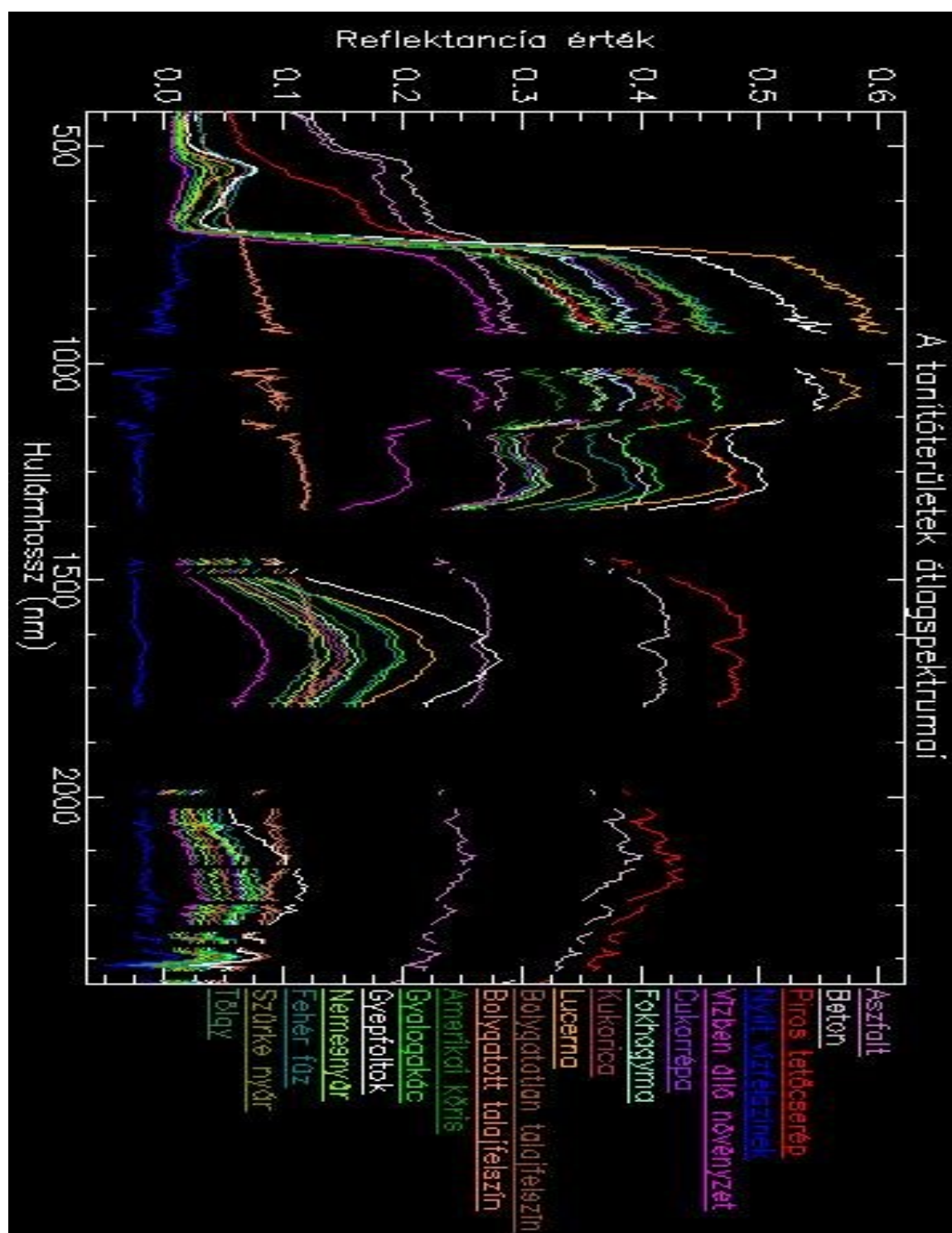
1. ábra: A doktori kutatás adatforrásai és fontosabb lépései.....	8
2. ábra: Egészséges és stressznek kitett vegetáció spektrumgörbéje, a reflektancia százalékában (forrás: Burai 2007)	16
3. ábra: A kriptobiotikus kéreg (BSC), a földutak és a szemiárid területek jellegzetes felszíntípusainak spektrumgörbéi, kontinuum-eltávolítás után (forrás: Weber et al. 2008).....	17
4. ábra: EO-1 Hyperion műholdfelvétel a Tisza-tó nyugati feléről (2005)	20
5. ábra: Tápai-rét és környéke az I. katonai felmérés térképszelvényén (SZTE Földrajzi Térképtár, EOVI vetületi rendszerbe korrigálva).....	22
6. ábra: A tájtérképezés mintaterülete, 1:10000-es EOVI térképszelvényen (1984, SZTE Földrajzi Térképtár).....	23
7. ábra: A 2010-es hiperspektrális légi felvételezés területe.....	25
8. ábra: A statisztikai tanszék (Rice) képfeldolgozó laboratóriuma	26
9. ábra: A hiperspektrális felvétel valószínűségi kivágata (a), a 981 nanométeres sáv szürkeárnyaltos képe (b), a 981 nm-es sáv Geary-féle C területi autokorrelációs térképe (c), az 1044 nm-es sáv szürkeárnyaltos képe (d) és az 1044 nm-es sáv Geary- féle C területi autokorrelációs térképe (e). A c, és az e, képeken világos szín jelzi a magasabb értékeket.....	31
10. ábra: A vizsgált osztályok kutatásomban használt szín- és betűkódjai.....	33
11. ábra: Az adatkomplexitás és a pontosság összefüggése (Landgrebe 2003, Hughes 1968).....	34
12. ábra: Az első főkomponens sáv szürkeárnyaltos képe.....	37
13. ábra: Képelemek a két főkomponenssáv spektrális terében	38
14. ábra: Kétféle T-próba szignifikancia (p) szintjei a szürkenyár és a gyalogakác spektrális mintáin, nem egyenlő varianciát feltételezve, a sávok száma alapján korrigálva	43
15. ábra: A szürkenyár és a gyalogakác Jeffries-Matusita módszerrel számított szeparabilitási értékei, spektrális sávonként	44
16. ábra: A vizsgált osztályok homogenitása, spektrális szögeltérés (rad) alapján	46
17. ábra: Gyalogakác a telepített nemesnyaras erdő cserjeszintjében (saját felvétel).....	50

18. ábra: Csongrád megye online kataszteri térképe	51
19. ábra: Az SAM módszerrel előállított felszínborítási térkép	55
20. ábra: A Maximum likelihood osztályozás eredménye (kivágat), a középső képen valószínűségi küszöb nélkül, az alsó képen 40%-os küszöbvel	59
21. ábra: Az optimális hipersík elhelyezkedése (Vapnik 1998).....	60
22. ábra: Az SVM osztályozás eredménye (kivágat).....	62
23. ábra: Az SOM klaszterek topológiai mátrixa	65
24. ábra: Az SOM irányított tanítása, a tanító- és tesztpontok százalékos pontosságának függvényében.....	66
25. ábra: Az SOM osztályozás eredménye (kivágat).....	68
26. ábra: Osztályozott képrészletek összehasonlítása: a, valós színes képkivágat b, az SVM térképe c, a Maximum likelihood térképe d, az SOM osztályozás eredménye	70
27. ábra: Mesterséges felszínek az SVM eredménytérképen	71
28. ábra: A talajfelszínek térképkivágata, SVM módszerrel osztályozva	73
29. ábra: A gyalogakác és a nemesnyár elkülönítése SAM osztályozással	74
30. ábra: Belvízelőöntés a mintaterületen, két nappal a felvételezés előtt (Készítette: Tobak Zsolt, 2010. szeptember 20.)	77
31. ábra: A belvízelőöntéssel érintett mintaterület valószínűségi képkivágata a hiperspektrális adaton	80
32. ábra: Binary encoding osztályozás eredménye a belvizes területen.....	83
33. ábra: Az SAM osztályozás eredménye a belvizes területen	84
34. ábra: Az ISODATA osztályozás eredménye: vízzel borított (kék), túlnedvesedett (lila), száraz (piros) talajfelszínek és a növényzet (zöld) osztálya	85
35. ábra: Az ISODATA eljárással előállított osztályok spektrumgörbéi	87
36. ábra: Az ISODATA osztályozás eredménye normalizált adaton	88
37. ábra: Az összehasonlító kivágat a hiperspektrális felvételen, valós színes megjelenítésben (bal felső ábra), a LiDAR domborzatmodellen (bal alsó ábra), az ISODATA osztályozás eredménytérképén (jobb felső ábra), a lefolyási irányokból számított összegyülekezési térkép (jobb alsó ábra).....	89
38. ábra: A városi felszínborítási osztályok box plot statisztikája spektrális szögeltérés felhasználásával (Tobak et al. 2013)	92
39. ábra: A városi felszínborítási osztályok a főkomponens sávok spektrális terében (Tobak et al. 2013)	93

40. ábra: A kertvárosi osztályok spektrális homogenitása szögeltérésük alapján (box plot diagramon)	95
41. ábra: A kertvárosi mintaterület valószínűségi kivágata (felső kép), illetve a Maximum likelihood osztályozás eredménye (alsó kép)	96
1. táblázat: A főkomponens sávok, és sajátérték rátájuk	36
2. táblázat: Az osztályok közötti Jeffries-Matusita (JM) és Transformed Divergence (TD) szeparabilitási értékek az első öt főkomponens sáv alapján	40
3. táblázat: Az osztályok közötti Jeffries-Matusita (JM) és Transformed Divergence (TD) szeparabilitási értékek az első két főkomponens sáv alapján	41
4. táblázat: Szórásértékek osztályonként az 50., 100., 150., 200. és 250. sávokban	45
5. táblázat: A szögeltérési (SAM) osztályozás pontosság-értékei	53
6. táblázat: Az SAM osztályozás tévesztési mátrixa	54
7. táblázat: A Maximum likelihood osztályozás tévesztési mátrixa	58
8. táblázat: A Maximum likelihood osztályozás pontosság értékei	58
9. táblázat: Az SVM osztályozás pontosság értékei	61
10. táblázat: Az SVM osztályozás tévesztési mátrixa	62
11. táblázat: Az SOM-hibrid osztályozás pontosság értékei	67
12. táblázat: Az SOM-hibrid osztályozás tévesztési mátrixa	67
13. táblázat: A lombkorona- és a cserjeszint osztályainak szeparabilitása	75
14. táblázat: A belvíztérképezési osztályok szeparabilitási értékei az első két főkomponens sáv alapján	81
15. táblázat: Az ISODATA eljárással előállított osztályok szeparabilitása	86
16. táblázat: Az ISODATA belvíztérképezési osztályok domborzati adatai (t. sz. f. m., méter).....	90
17. táblázat: A városi felszínborítási osztályok spektrális elkülöníthetősége (Tobak et al. 2012a).....	92
18. táblázat: A kertvárosi tanítóterületek Jeffries-Matusita (JM) és Transformed Divergence (TD) szeparabilitási értékei	94

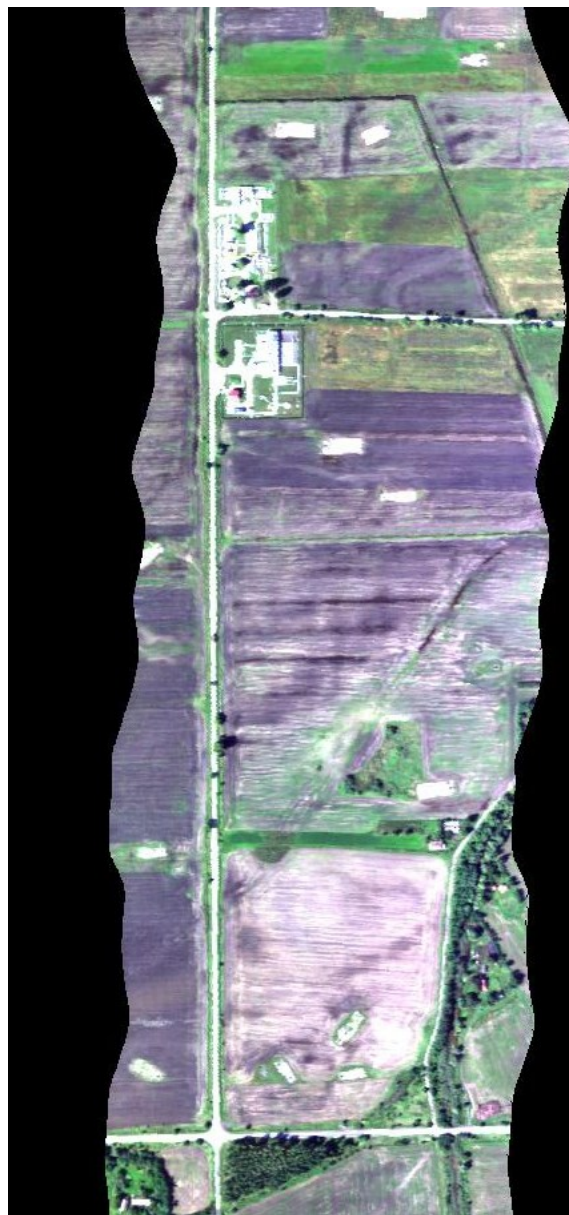
Mellékletek

1. melléklet

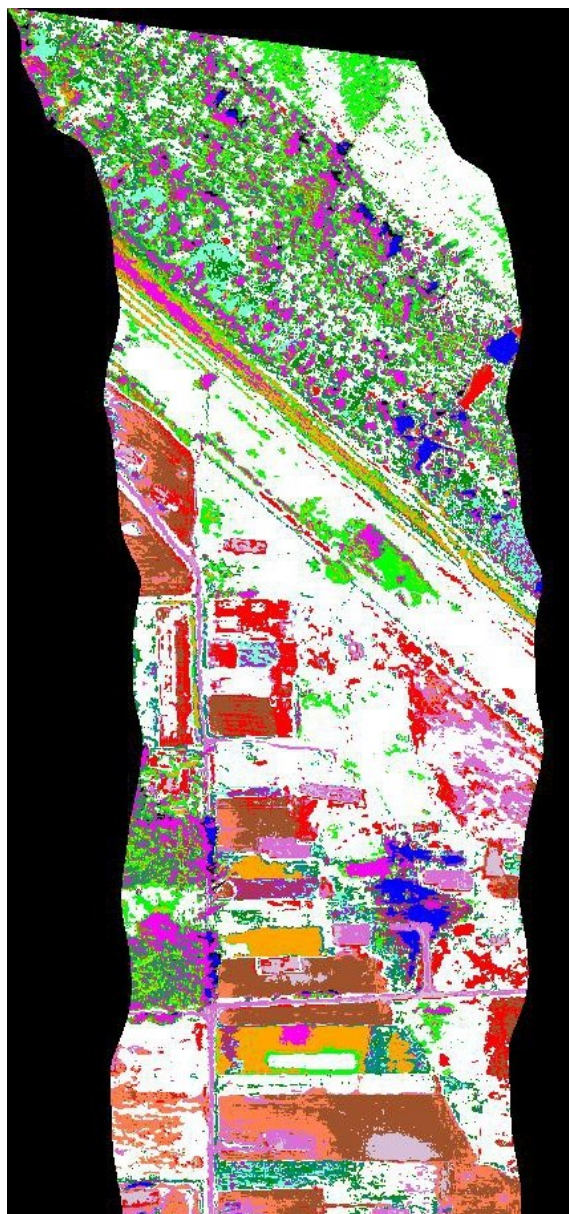
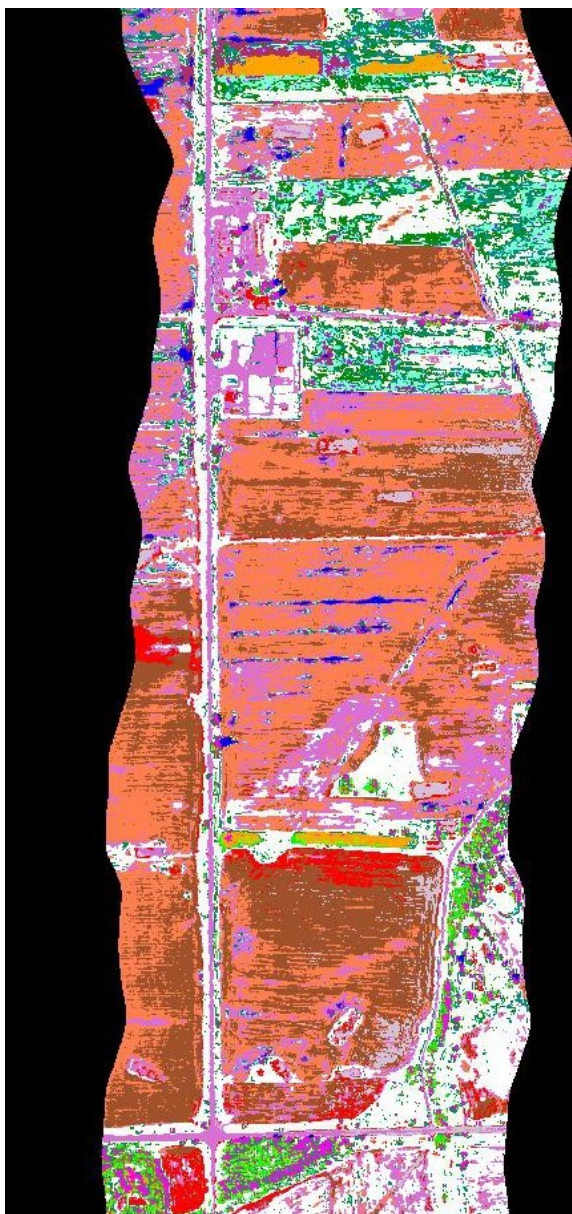


A felhasznált tanítóterületek reflektanciagörbéi

2. melléklet: Az osztályozási térképek



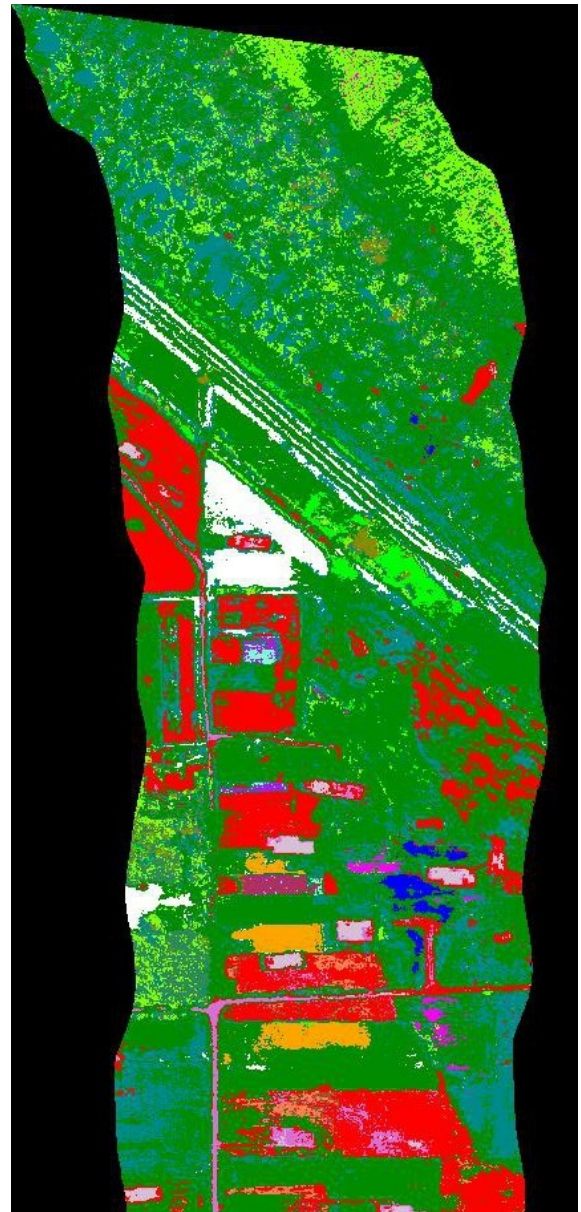
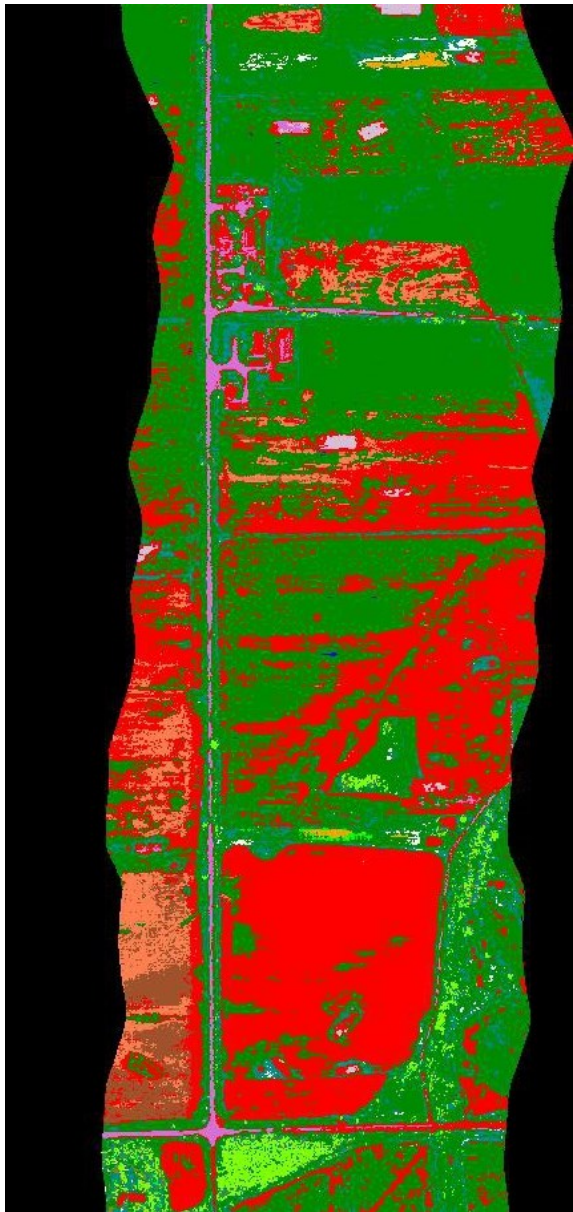
Az eredeti hiperspektrális felvétel, valószínűs megjelenítésben (baloldalon: a terület keleti fele, elforgatva, jobboldalon: a terület nyugati fele, elforgatva)



Osztályok

 Amerikai kőris (A)	 Gyalogakác (G)	 Vízben álló növényzet (U)
 Aszfalt (B)	 Gyepfoltok (H)	 Nyílt vízfelület (V)
 Beton (C)	 Nemesnyaras (J)	 Lucerna (W)
 Bolygatatlan talajfelszín (D)	 Fehér fűz (K)	 Kukorica (X)
 Bolygatott talajfelszín (E)	 Szürke nyár (M)	 Fokhagyma (Y)
 Piros tetőcserép (F)	 Tölgy (N)	 Cukorrépa (Z)

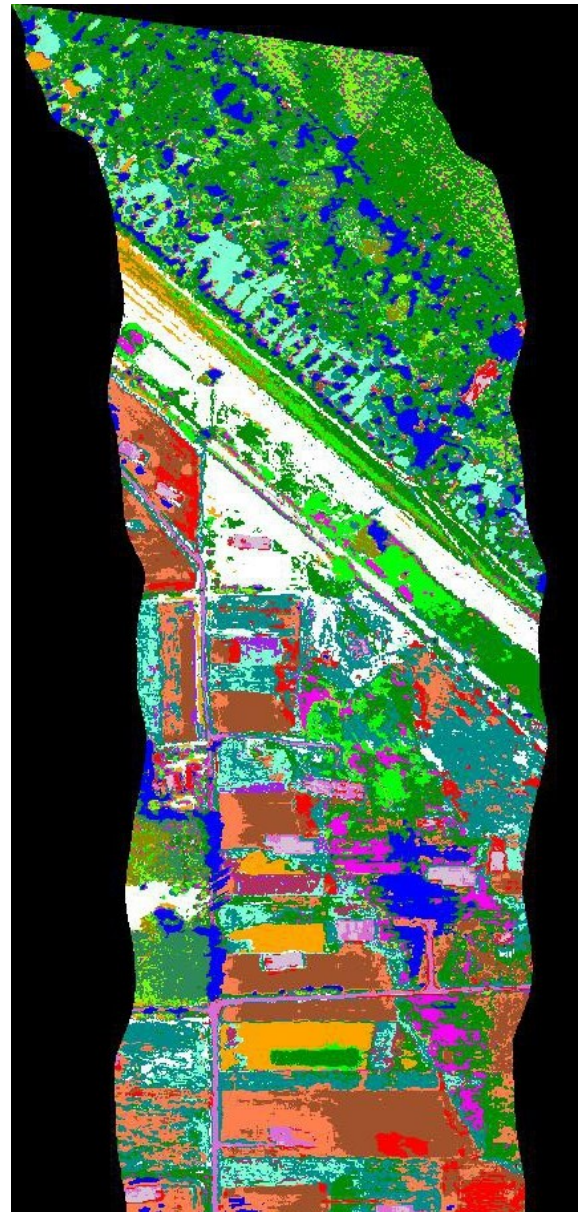
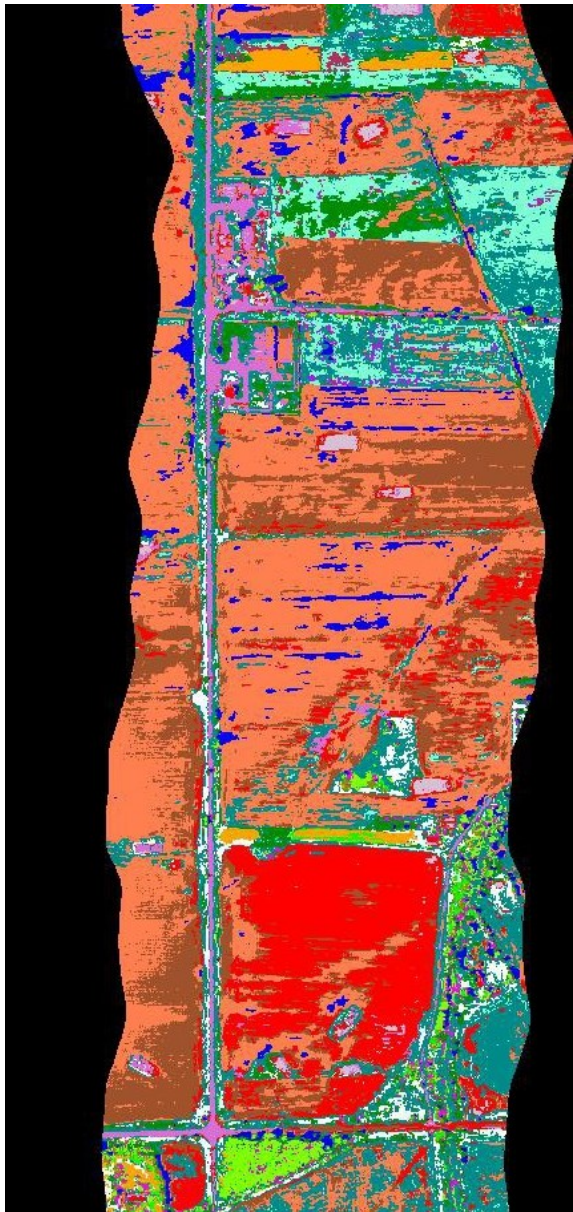
Az SAM osztályozás eredménye (baloldalon: a terület keleti fele, elforgatva, jobboldalon: a terület nyugati fele, elforgatva)



Osztályok

 Amerikai kőris (A)	 Gyalogakác (G)	 Vízben álló növényzet (U)
 Aszfalt (B)	 Gyepfoltok (H)	 Nyílt vízfelület (V)
 Beton (C)	 Nemesnyaras (J)	 Lucerna (W)
 Bolygatatlan talajfelszín (D)	 Fehér fűz (K)	 Kukorica (X)
 Bolygatott talajfelszín (E)	 Szürke nyár (M)	 Fokhagyma (Y)
 Piros tetőcserép (F)	 Tölgy (N)	 Cukorrépa (Z)

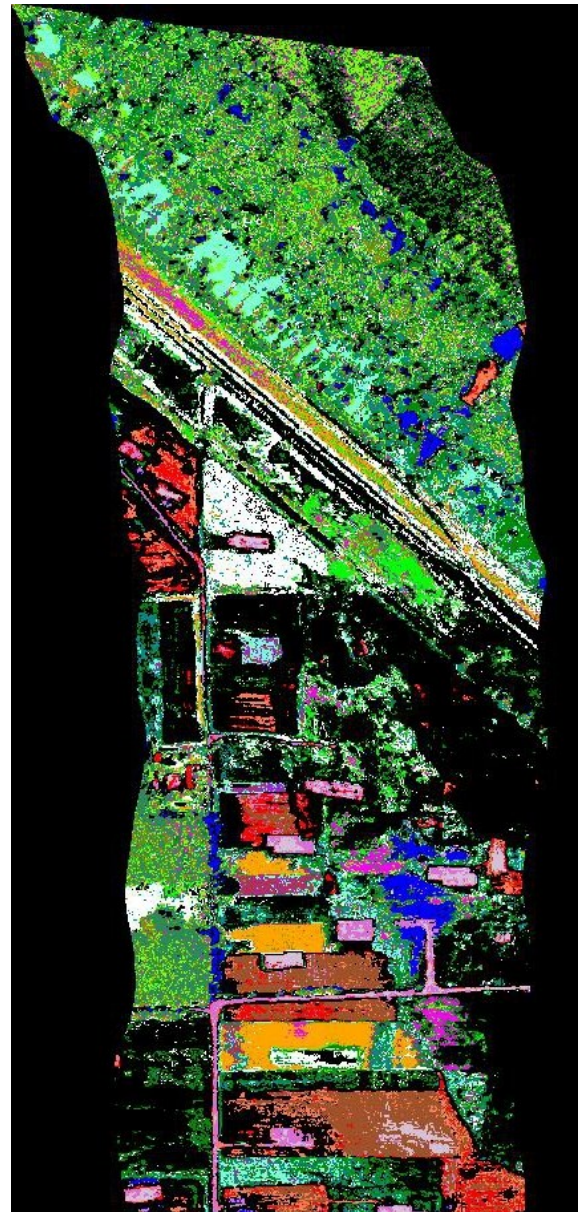
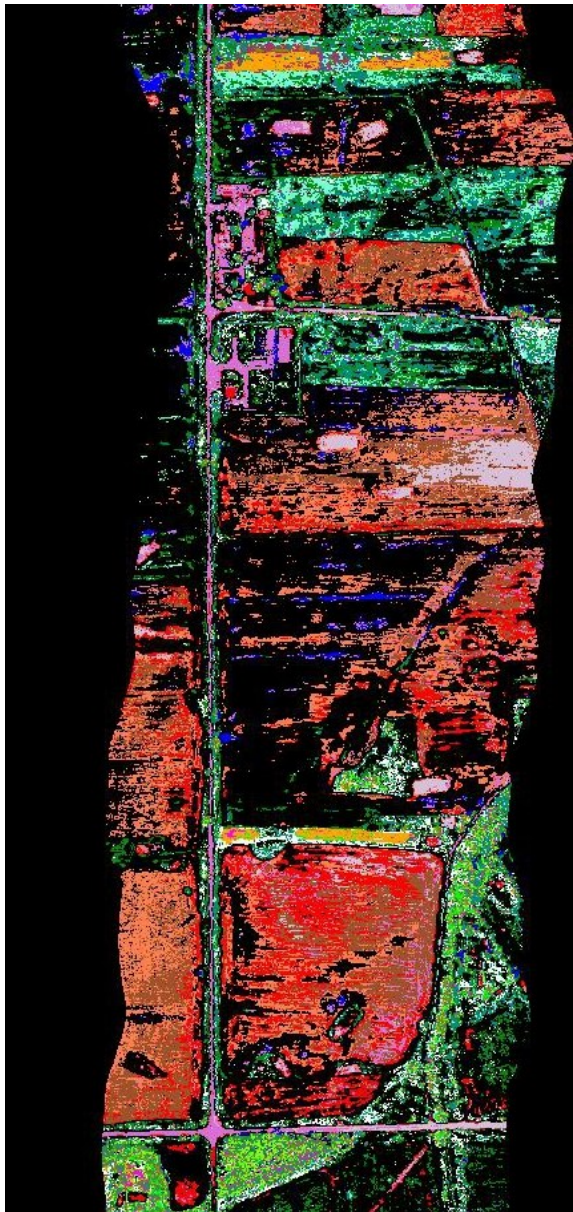
A Maximum likelihood osztályozás eredménye (baloldalon: a terület keleti fele, elforgatva, jobboldalon: a terület nyugati fele, elforgatva)



Osztályok

 Amerikai kőris (A)	 Gyalogakác (G)	 Vízben álló növényzet (U)
 Aszfalt (B)	 Gyepfoltok (H)	 Nyílt vízfelület (V)
 Beton (C)	 Nemesnyaras (J)	 Lucerna (W)
 Bolygatatlan talajfelszín (D)	 Fehér fűz (K)	 Kukorica (X)
 Bolygatott talajfelszín (E)	 Szürke nyár (M)	 Fokhagyma (Y)
 Piros tetőcserép (F)	 Tölgy (N)	 Cukorrépa (Z)

Az SVM osztályozás eredménye (baloldalon: a terület keleti fele, elforgatva, jobboldalon: a terület nyugati fele, elforgatva)



Osztályok

 Amerikai kőris (A)	 Gyalogakác (G)	 Vízben álló növényzet (U)
 Aszfalt (B)	 Gyepfoltok (H)	 Nyílt vízfelület (V)
 Beton (C)	 Nemesnyaras (J)	 Lucerna (W)
 Bolygatatlan talajfelszín (D)	 Fehér fűz (K)	 Kukorica (X)
 Bolygatott talajfelszín (E)	 Szürke nyár (M)	 Fokhagyma (Y)
 Piros tetőcserép (F)	 Tölgy (N)	 Cukorrépa (Z)

A Self-organizing map (SOM) osztályozás eredménye (baloldalon: a terület keleti fele, elforgatva, jobboldalon: a terület nyugati fele, elforgatva)

3. melléklet: Zajos sávok listája

Táblázatomban 0 jelöli a zajosnak, 1 a zajmentesnek ítélt adatsávokat, melyeket hullámhosszukkal jelöltem.

Hullámhossz (nanométer)	Zajos (0) / zajmentes (1)	Hullámhossz (nanométer)	Zajos (0) / zajmentes (1)	Hullámhossz (nanométer)	Zajos (0) / zajmentes (1)
401	0	489	1	580	1
405	0	493	1	584	1
409	0	498	1	589	1
414	0	502	1	594	1
418	1	507	1	598	1
422	1	511	1	603	1
427	1	516	1	608	1
431	1	520	1	612	1
435	1	525	1	617	1
440	1	529	1	622	1
444	1	534	1	626	1
449	1	538	1	631	1
453	1	543	1	635	1
458	1	547	1	640	1
462	1	552	1	645	1
466	1	556	1	649	1
471	1	561	1	654	1
475	1	566	1	659	1
480	1	570	1	663	1
484	1	575	1	668	1

Hullámhossz (nanométer)	Zajos (0) / zajmentes (1)	Hullámhossz (nanométer)	Zajos (0) / zajmentes (1)	Hullámhossz (nanométer)	Zajos (0) / zajmentes (1)
673	1	833	1	1006	0
677	1	838	1	1013	1
682	1	842	1	1019	1
687	1	847	1	1025	1
691	1	852	1	1031	1
696	1	857	1	1038	1
701	1	861	1	1044	1
705	1	866	1	1050	1
710	1	871	1	1057	1
715	1	876	1	1063	1
719	1	880	1	1069	1
724	1	885	1	1075	1
729	1	890	1	1082	1
733	1	895	1	1088	1
738	1	899	1	1094	1
743	1	904	1	1101	1
747	1	909	1	1107	1
752	1	914	1	1113	0
757	1	919	1	1119	0
761	1	923	1	1126	0
766	1	928	1	1132	1
771	1	933	0	1138	1
776	1	938	0	1144	1
781	1	942	0	1151	1
785	1	947	0	1157	0
790	1	952	0	1163	1
795	1	957	0	1170	1
800	1	961	0	1176	1
804	1	969	0	1182	1
809	1	975	0	1188	1
814	1	981	0	1195	1
819	1	987	0	1201	1
823	1	994	0	1207	1
828	1	1000	0	1214	1

Hullámhossz (nanométer)	Zajos (0) / zajmentes (1)	Hullámhossz (nanométer)	Zajos (0) / zajmentes (1)	Hullámhossz (nanométer)	Zajos (0) / zajmentes (1)
1220	1	1433	0	1647	1
1226	1	1440	0	1653	1
1232	1	1446	0	1660	1
1239	1	1452	1	1666	1
1245	1	1459	1	1672	1
1251	1	1465	1	1678	1
1258	1	1471	0	1685	1
1264	1	1477	1	1691	1
1270	1	1484	1	1697	1
1276	1	1490	0	1703	1
1283	1	1496	1	1710	1
1289	1	1502	1	1716	1
1295	1	1509	1	1722	1
1301	1	1515	1	1729	1
1308	1	1521	1	1735	1
1314	1	1528	1	1741	1
1320	1	1534	1	1747	1
1327	1	1540	1	1754	1
1333	1	1546	1	1760	1
1339	1	1553	1	1766	1
1345	0	1559	1	1773	1
1352	0	1565	1	1779	1
1358	0	1572	1	1785	1
1364	0	1578	1	1791	1
1371	0	1584	1	1798	0
1377	0	1590	1	1804	0
1383	0	1597	1	1810	0
1389	0	1603	1	1817	0
1396	0	1609	1	1823	0
1402	0	1616	1	1829	0
1408	0	1622	1	1835	0
1415	0	1628	1	1842	0
1421	0	1634	1	1848	0
1427	0	1641	1	1854	0

Hullámhossz (nanométer)	Zajos (0) / zajmentes (1)	Hullámhossz (nanométer)	Zajos (0) / zajmentes (1)	Hullámhossz (nanométer)	Zajos (0) / zajmentes (1)
1860	0	2061	1	2262	1
1867	0	2068	1	2269	1
1873	0	2074	1	2275	1
1879	0	2080	1	2281	1
1886	0	2087	1	2288	1
1892	0	2093	1	2294	1
1898	0	2099	1	2300	0
1904	0	2105	1	2306	0
1911	0	2112	1	2313	1
1917	0	2118	1	2319	1
1923	0	2124	1	2325	1
1930	0	2131	1	2332	1
1936	0	2137	1	2338	1
1942	0	2143	1	2344	0
1948	0	2149	1	2350	1
1955	0	2156	1	2357	1
1961	0	2162	0	2363	1
1967	0	2168	1	2369	1
1974	1	2175	1	2376	1
1980	0	2181	1	2382	1
1986	1	2187	1	2388	1
1992	1	2193	1	2394	1
1999	0	2200	1	2401	1
2005	0	2206	1	2407	0
2011	0	2212	1	2413	0
2018	0	2219	1	2419	0
2024	0	2225	1	2426	1
2030	1	2231	1	2432	1
2036	1	2237	1	2438	0
2043	1	2244	0	2445	0
2049	1	2250	1	2451	0
2055	1	2256	1	Összesen	273

4. melléklet: Az osztályok szögeltérési statisztikái

Osztály	Min	Max	Átlag	Szórás	Q1	Q3
Amerikai kőris (A)	0,032854	0,23401	0,101267	0,045816	0,0715	0,1209
Aszfalt (B)	0,021227	0,111132	0,05047	0,02317	0,0319	0,063
Beton (C)	0,040225	0,112709	0,069486	0,015848	0,0563	0,081
Bolygatatlan talaj (D)	0,030586	0,07525	0,048274	0,010598	0,0387	0,0561
Bolygatott talaj (E)	0,036106	0,103921	0,056624	0,014981	0,0469	0,0651
Cserép (F)	0,037784	0,346649	0,081412	0,046149	0,0596	0,092
Gyalogakác (G)	0,014379	0,117533	0,0493	0,02635	0,0294	0,0688
Gyep (H)	0,01573	0,092425	0,042526	0,020449	0,0259	0,0572
Nedves talaj (I)	0,070743	1,11758	0,332941	0,323738	0,1275	0,3738
Nemesnyár (J)	0,020328	0,25953	0,083253	0,052841	0,0474	0,1042
Fehér fűz (K)	0,024976	0,213395	0,083123	0,047897	0,044	0,1019
Száraz talaj (L)	0,019884	0,073084	0,0358	0,012316	0,0278	0,0385
Szürke nyár (M)	0,021193	0,264043	0,061783	0,040577	0,0368	0,0752
Tölgy (N)	0,023872	0,160743	0,075146	0,035603	0,0451	0,0979
Vizben álló növény (u)	0,021739	0,190227	0,092182	0,044189	0,0594	0,1129
Nyílt víz (v)	0,135727	1,008708	0,299131	0,195572	0,183	0,2989
Lucerna (w)	0,012374	0,026489	0,019025	0,00393	0,0153	0,0218
Kukorica (x)	0,013403	0,049419	0,02548	0,008156	0,0196	0,0289
Fokhagyma (y)	0,016462	0,093094	0,042282	0,021658	0,0243	0,0492
Cukorrépa (z)	0,018806	0,065006	0,037102	0,014721	0,0233	0,0509

5. Melléklet: Spektrális szeparabilitás sávonként

Az alábbi táblázat a szürkenyár és a gyalogakác tanítópontjainak elkülöníthetőségét írja le az egyes spektrális adatsávokat. A számításokat Jeffries-Matusita (J-M) indexszel, illetve kétmintás T-próbával, a szignifikancia (p) értékek mérésével végeztem el.

Hullámhossz (nanométer)	J-M index	T-próba szignifikancia	Hullámhossz (nanométer)	J-M index	T-próba szignifikancia
414	0,285	1	502	0,68	5,52E-10
418	0,369	1	507	0,693	1,75E-10
422	0,446	1	511	0,726	1,44E-11
427	0,458	1	516	0,749	4,30E-14
431	0,438	1	520	0,728	9,62E-13
435	0,413	1	525	0,71	4,62E-13
440	0,516	0,096141	529	0,695	2,69E-12
444	0,546	0,443542	534	0,67	3,79E-12
449	0,492	0,056389	538	0,675	1,23E-12
453	0,563	0,009971	543	0,682	1,62E-12
458	0,554	0,001041	547	0,668	1,32E-12
462	0,573	0,00084	552	0,68	4,18E-13
466	0,609	4,07E-05	556	0,693	4,22E-14
471	0,594	0,0002	561	0,731	1,73E-14
475	0,612	4,13E-06	566	0,771	4,07E-16
480	0,622	2,26E-06	570	0,831	5,24E-18
484	0,606	5,35E-06	575	0,853	1,45E-18
489	0,625	1,96E-05	580	0,858	3,48E-18
493	0,625	4,69E-08	584	0,878	4,99E-19
498	0,651	1,08E-07	589	0,901	1,21E-19

Hullámhossz (nanométer)	J-M index	T-próba szignifikancia	Hullámhossz (nanométer)	J-M index	T-próba szignifikancia
594	0,914	2,51E-19	752	0,475	4,58E-06
598	0,925	1,79E-20	757	0,488	1,00E-05
603	0,939	3,79E-20	761	0,552	1,03E-08
608	0,928	1,13E-19	766	0,505	1,09E-07
612	0,93	5,24E-20	771	0,503	8,38E-07
617	0,937	5,71E-20	776	0,495	1,06E-06
622	0,933	3,43E-20	781	0,493	9,87E-07
626	0,937	8,46E-20	785	0,51	1,53E-07
631	0,943	1,09E-19	790	0,51	2,50E-07
635	0,956	2,92E-20	795	0,505	4,95E-07
640	0,937	2,87E-20	800	0,512	2,69E-07
645	0,961	2,94E-20	804	0,516	5,81E-08
649	0,952	2,46E-20	809	0,517	1,19E-07
654	0,918	6,71E-19	814	0,523	6,60E-08
659	0,905	3,59E-18	819	0,514	2,28E-07
663	0,864	5,58E-16	823	0,515	4,87E-08
668	0,817	1,12E-14	828	0,534	2,81E-08
673	0,776	3,37E-13	833	0,523	1,58E-08
677	0,777	4,87E-15	838	0,518	3,87E-08
682	0,775	1,40E-11	842	0,543	1,52E-09
687	0,758	1,22E-12	847	0,533	4,42E-09
691	0,825	3,25E-15	852	0,538	2,49E-09
696	0,845	7,18E-17	857	0,54	9,83E-09
701	0,788	9,19E-16	861	0,539	4,44E-09
705	0,684	1,72E-13	866	0,534	4,26E-09
710	0,555	2,02E-09	871	0,539	2,29E-09
715	0,429	2,71E-05	876	0,54	2,91E-09
719	0,291	0,247703	880	0,55	2,51E-10
724	0,272	1	885	0,55	4,53E-10
729	0,284	1	890	0,554	1,55E-10
733	0,322	1	895	0,558	6,32E-09
738	0,381	0,047132	899	0,591	4,95E-12
743	0,429	0,000505	904	0,567	1,75E-10
747	0,454	6,87E-05	909	0,605	4,31E-12

Hullámhossz (nanométer)	J-M index	T-próba szignifikancia	Hullámhossz (nanométer)	J-M index	T-próba szignifikancia
914	0,58	6,86E-11	1484	0,917	2,57E-18
919	0,596	5,63E-13	1496	0,811	5,81E-13
923	0,596	4,62E-12	1502	0,797	1,32E-12
987	0,802	1,94E-25	1515	0,791	1,18E-12
994	0,802	9,03E-26	1521	0,8	7,21E-13
1000	0,789	1,74E-25	1528	0,801	3,46E-13
1006	0,772	7,03E-26	1534	0,8	5,04E-13
1019	0,763	2,24E-23	1540	0,8	2,03E-13
1025	0,733	2,88E-21	1546	0,804	8,48E-14
1031	0,751	3,24E-21	1553	0,803	4,66E-14
1038	0,756	6,36E-21	1559	0,808	4,05E-14
1044	0,753	6,99E-21	1565	0,811	1,45E-14
1050	0,751	4,14E-21	1572	0,816	9,86E-15
1057	0,739	1,55E-19	1578	0,815	7,63E-15
1063	0,742	1,15E-19	1584	0,81	7,91E-15
1069	0,742	7,94E-20	1590	0,811	1,07E-14
1207	0,793	1,70E-18	1597	0,807	1,36E-14
1214	0,797	8,27E-19	1603	0,812	5,30E-15
1220	0,788	3,40E-18	1609	0,806	7,83E-15
1226	0,776	1,05E-17	1616	0,804	1,04E-14
1232	0,768	2,84E-17	1622	0,802	1,34E-14
1239	0,759	5,19E-17	1628	0,797	1,41E-14
1245	0,755	1,61E-16	1634	0,799	1,04E-14
1251	0,75	1,22E-16	1641	0,795	7,56E-15
1258	0,746	8,77E-17	1647	0,798	8,63E-15
1264	0,752	4,33E-17	1653	0,8	6,53E-15
1276	0,739	4,76E-16	1660	0,798	7,07E-15
1283	0,739	5,56E-16	1666	0,804	4,67E-15
1301	0,738	3,84E-16	1672	0,805	4,69E-15
1308	0,745	3,34E-16	1678	0,803	4,48E-15
1314	0,75	1,32E-16	1685	0,805	3,52E-15
1320	0,758	1,26E-16	1691	0,808	5,98E-15
1327	0,764	2,19E-17	1697	0,811	1,44E-15
1477	0,914	1,28E-19	1703	0,814	1,01E-15

Hullámhossz (nanométer)	J-M index	T-próba szignifikancia	Hullámhossz (nanométer)	J-M index	T-próba szignifikancia
1710	0,819	8,81E-16	2149	0,759	8,82E-11
1716	0,828	2,65E-16	2156	0,762	1,20E-11
1722	0,836	7,35E-17	2175	0,757	1,34E-11
1729	0,851	1,90E-17	2181	0,755	1,76E-10
1735	0,85	1,03E-17	2187	0,764	3,86E-11
1741	0,844	5,90E-17	2193	0,767	2,14E-11
1747	0,83	2,28E-16	2200	0,778	4,59E-12
1754	0,829	1,47E-16	2206	0,765	1,37E-11
1760	0,845	3,84E-17	2212	0,766	8,37E-12
1766	0,837	7,44E-17	2219	0,782	1,80E-12
1779	0,841	6,83E-17	2225	0,786	1,75E-12
1785	0,843	1,09E-17	2231	0,781	1,26E-12
2011	0,695	3,95E-18	2237	0,788	9,86E-13
2018	0,718	1,40E-08	2244	0,799	2,07E-12
2024	0,658	6,10E-05	2250	0,813	3,17E-14
2030	0,587	0,001414	2256	0,821	7,54E-15
2036	0,62	0,001808	2262	0,836	9,92E-15
2043	0,641	9,59E-05	2275	0,834	3,43E-15
2055	0,696	1,68E-07	2288	0,879	4,37E-17
2061	0,73	4,39E-08	2294	0,882	1,68E-19
2068	0,751	9,30E-10	2313	0,876	1,04E-19
2080	0,733	1,76E-08	2319	0,875	8,10E-16
2087	0,722	5,08E-09	2332	0,832	1,38E-14
2093	0,744	1,70E-09	2338	0,824	4,99E-13
2099	0,757	2,91E-10	2357	0,804	1,47E-14
2105	0,757	8,45E-10	2363	0,808	3,02E-13
2112	0,768	1,03E-09	2369	0,829	4,03E-17
2118	0,759	7,75E-11	2376	0,809	1,16E-16
2124	0,753	7,57E-11	2382	0,841	4,06E-17
2131	0,772	4,36E-11	2388	0,852	3,49E-16
2137	0,801	3,89E-11	2394	0,783	8,30E-14
2143	0,761	7,99E-11	2401	0,77	3,38E-14