

Magyarország Potenciális Vegetációjának Modellezése

OTKA PD 83522 pályázat zárójelentése

Somodi Imelda

MTA Ökológiai Kutatóközpont, Ökológiai és Botanikai Intézet

2015.10.30

1. Fogalmi alapozás

A potenciális vegetáció fogalma hosszú múltra tekinthet vissza (Tüxen 1956) és az eltelt időben egyrészt széles körben alkalmazottá vált, másrészt számos, az eredetitől eltérő felfogás terjedt el. Ezért – az eredetileg tervezett munkaterven felül - mindenekelőtt az aktuális potenciális vegetáció (Potential Natural Vegetation, PNV) fogalmának tisztázására volt szükség, hogy a későbbi publikációkban erre már hivatkozassanak. A PNV fogalmat ugyanis gyakran nem választották el a prehumán (rekonstruált) vegetáció foglamától, ami számos félreértésre és megalapozatlan elvárásra adott okot (pl. Chiarucci és mtsai. 2010). Az eredeti definícióhoz visszatérve azonban egyértelmű, hogy a PNV az aktuális viszonyokra vonatkozik, tehát azt a vegetációt jelenti, amely számára az aktuális környezeti viszonyok megfelelőek, amelyek eltérhetnek, sőt európai viszonylatban általában eltérnek a prehumán viszonyoktól. A kettő közötti különbségtétel tette azt is lehetővé, hogy a PNV fogalom a Pannon vegetáció szabályosságainak leírásánál is hasznosítható legyen (Fekete és mtsai. 2014). A PNV-nek megfelelő aktuális vegetációt a szabályosságok körébe lehet sorolni, míg deviáció ott fordul elő, ahol az aktuális vegetáció például történeti okból, a vegetációs inertia következtében eltér a PNV-től. Azt is megállapítottuk, hogy az extrazonalitás indikátora is tud lenni a PNV-től való különbözőség (Fekete és mtsai 2011).

Az elméleti háttér tisztázása mellett társszerzőimmal javaslatot tettünk a PNV felfogás operatívabbá tételére (Somodi és mtsai. 2012). Bevezettük a többretegű potenciális vegetáció (a továbbiakban Multiple Potential Natural Vegetation, MPNV) fogalmát, amely lehetővé teszi, hogy egy hely potenciális vegetációját nem egyetlen típusként, hanem típusok valószínűségi eloszlásaként értelmezzünk.

2. Magyarország MPNV modellje

A projekt központi eleme a fenti elméleti keret megvalósítása a gyakorlatban: Magyarország aktuális potenciális vegetációmodelljének elkészítése. Bár elméleti alapozó cikkünk nyomán (azt idézve) számos új, immár operatív, formalizált és statisztikai alapú PNV becslés történt (Hemsing és Bryn 2012, Attore és mtsai. 2014, Reger és mtsai. 2014), eloszlás-alapú MPNV becslésre nem történt kísérlet. Valószínűsíthetően azért is, mert a potenciális vegetáció modellezése speciális kihívásokkal szembesíti a kutatókat. Az egyik ilyen a PNV modellek jóságának értékelése, a másik a párhuzamos valószínűségbecslések összehasonlíthatóvá tétele. Ezért a következő célokat fogalmaztuk meg:

- Az MPNV modellsorozat elkészítése
- Új PNV-értékelési módszerek kidolgozása
- A valószínűségi rétegek összehasonlíthatóvá tétele

Módszerek

A modellezés alapja a még megmaradt természetes élőhelyek és élettelen környezetük között megfigyelt összefüggések statisztikai formalizálása. Az élőhelyi információkat a MÉTA adatbázisból nyertük. Olyan élőhelyeket választottunk, amelyek természetesek és a növényzet természetes fejlődésének végső állomásait képviselik, a tágabb értelemben vett klímax vegetációhoz sorolhatók (1. táblázat). Ezek egy részénél a térképezés óta eltelt tapasztalatok alapján kiderült, hogy elkülönítésük nem volt stabil, függött a térképező személyétől (tapasztalatától, szakterületétől), ezért az ÁNÉR osztályozáshoz képest összevonásokra került sor (1. táblázat).

1. táblázat. A modellezett élőhelyek. Elsődlegesen a MÉTA térképezéskor használt ÁNÉR élőhelyek közül választottuk ki a természetes, stabil élőhelyeket. A térképezés óta felgyűlt tapasztalatok alapján néhányat összevontunk. Ezeknél mindegyik eredeti élőhelykódot feltüntettem.

Élőhelyek		Modelljóság		
		AUC	AUCD	MPD
B1a	Nem tűzegképző nádasok, gyékényesek és tavikákások	0.890	0.050	0.063
B1b	Úszólápok, tűzeges nádasok és télisásosok	0.919	0.064	0.222
B4	Lápi zsombékosok, zsombék-semlyék komplexek	0.857	0.089	0.101
B6	Zsiókás, kötő kákás és nádas szikes vizű mocsarak	0.946	0.030	0.103
F1a	Ürmöspuszták	0.969	0.009	0.030
F2	Szikes rétek	0.973	0.017	0.058

F4	Üde mézpázsitos szikfokok	0.958	0.026	0.119
F5	Padkás szikések, szikes tavak iszap- és vakszik növényzete	0.949	0.025	0.078
G1	Nyílt homokpusztagyepek	0.968	0.017	0.091
G2	Mészkedvelő nyílt sziklagyepek	0.955	0.027	0.147
G3	Nyílt szilikátsziklagyepek és törmeléklejtők	0.952	0.032	0.139
H1	Zárt sziklagyepek	0.908	0.089	0.346
H2	Felnyíló, mészkedvelő lejtő- és törmelékgyepek	0.978	0.019	0.198
H3a	Köves talajú lejtősztyepek	0.950	0.027	0.115
H4	Erdőssztyeprétek, félszáraz irtásrétek, száraz magaskórósok	0.931	0.047	0.151
H5a	Lőszgyepek, kötött talajú sztyeprétek	0.928	0.039	0.089
H5b	Homoki sztyeprétek	0.971	0.017	0.090
J1a	Fűzlápok	0.905	0.064	0.137
J2	Láp- és mocsárerdők	0.932	0.039	0.116
J3_J4	Folyómenti bokorfüzesek - Fűz-nyár ártéri erdők	0.985	0.011	0.093
J5	Égerligetek	0.952	0.017	0.057
J6	Keményfás ártéri erdők	0.952	0.035	0.156
K1a_K2_K7b	Gyertyános-tölgyesek	0.946	0.027	0.054
K5_K7a	Bükkösök	0.954	0.008	0.033
L1_M1	Mész- és melegkedvelő tölgyesek - Molyhos tölgyes bokorerdők	0.958	0.015	0.072
L2a_L2b	Cseres-tölgyesek	0.950	0.029	0.086
L2x	Hegylábi zárt erdőssztyep tölgyesek	0.970	0.011	0.089
L4a_L4b	Mészkerülő tölgyesek	0.967	0.024	0.198
L5	Alföldi zárt kocsányos tölgyesek	0.967	0.030	0.254
LY1	Szurdokerdők	0.962	0.022	0.222
LY2	Törmeléklejtő-erdők	0.943	0.021	0.115
LY3	Bükkös sziklaerdők	0.970	0.028	0.282
LY4	Tölgyes jellegű sziklaerdők és tetőerdők	0.959	0.019	0.138
M2	Nyílt löszölgyesek	0.820	0.179	0.078
M3	Nyílt sziki tölgyesek	0.919	0.081	0.342
M4	Nyílt homoki tölgyesek	0.940	0.059	0.354
M5	Homoki borókás-nyárasok	0.989	0.010	0.185
M6	Sztyepcserjések	0.888	0.098	0.178
M7	Sziklai cserjések	0.914	0.046	0.201
N13	Mészkerülő lombelegyes fenyvesek	0.998	0.001	0.145

Az élettelen környezetet különböző háttérváltozók jelenítik meg (2. táblázat). A vegetáció és az élettelen környezet viszonyát Gradient Boosting Modell (GBM) segítségével formalizáltuk. Mivel az utóbbi években fény derült arra, hogy ez a technika sem tolerálja a háttérváltozók közötti magas korrelációt ezért olyan kiinduló adatbázist építettünk fel, amelyben a változók közötti páros korrelációk 0.8 és a multikollinearitás értéke a variance inflation factor (VIF) szerint 30 alatti. Így ugyan az összes, a WORDCLIM által javasolt bioklimatikus változót kiszámoltuk, de csak a 2. táblázatban szereplők kerülhettek be a kiindulási adatbázisba.

2. táblázat. A modellekben felhasznált környezeti háttérváltozók.

Rövidítés	Leírás	Forrás
<i>Bioklimatikus változók</i>		
BIO2	Átlagos napi hőingás	Leskálázott WORDCLIM adatbázis (Hijmans et al., 2005)
BIO4	A hőmérséklet évszakossága	
BIO8	A legnedvesebb negyedév (nyár) átlaghőmérséklete	
BIO11	A leghidegebb negyedév (tél) átlaghőmérséklete	Magyarország Meteorológiai Atlasza (OMSz, 2001)
BIO16	A legnedvesebb negyedév csapadékmennyisége	
BIO19	A leghidegebb negyedév csapadékmennyisége	
<i>Talajadatok</i>		
Sbr	Alapkőzet	AGROTOPO adatbázis (Várallyay et al., 1994)
Spt	Fizikai talajtípus	
Swb	Vízgazdálkodás	
Sph	Kémhatás és mészállapot	
Soc	Szervesanyag készlet	
Shd	Termőréteg vastagság	
<i>Hydrology</i>		
Dla	A legközelebbi tótól való távolság	Digitalizált vektoros adat
Dri	A legközelebbi folyótól való távolság	
Dst	A legközelebbi pataktól való távolság	
Dwa	A legközelebbi víztest távolsága (csatornákat is ideértve)	

A modellépítéshez nem használtuk fel azokat a MÉTA hatszögeket, amelyek egyáltalán nem tartalmaztak természetes élőhelyet, hogy csökkentsük azon hiányinformációk befolyását, amelyek az emberi behatásnak és nem az élőhely környezeti preferenciájának következményei.

Az így szűrt adatbázis 90000-es nagyságrendű adateleme lehetővé tette, hogy minden élőhelynél külön-külön a jelenlét/hiány arányt megtartva tanuló és tesztadatokra bontsuk az adatbázist.

Alternatív modellértékelési lehetőségek

A PNV modellek értékelésének legnagyobb kihívása az általában hibának tekintett, a predikció és a megfigyelés közötti eltérések kezelése. Az értékelésben ugyanis a legtöbbször az un. kontingencia-táblára (3. táblázat) hagyatkoznak és azt a modellt tartják a legjobbnak, amelynél a hibás becslések alacsonyak. Csakhogy a PNV modellezés éppen arra irányul, hogy megbecsüljük

adott, éppen akár vegetációval nem rendelkező helyen a lehetséges eloszlást, tehát a hibás jelenlétebecslések valójában a PNV-becslés célját jelentik. Kevésbé egyértelmű, de a hibás hiánybecslések is lehetnek információ hordozói. Ugyanis a megfigyelt vegetáció, különösen ha fás szárú vegetációról van szó, fennmaradhat számára már nem ideális körülmények között is. Az ilyen állományokra mutat rá a hibás hiánybecslések egy része, amellet, hogy természetesen egy hibakomponenst is hordoz.

Ennek a paradoxonnak a feloldására vezettünk be két új mértéket. Az alapelv mindkét esetben azon nyugszik, hogy a modellek mindig jobban teljesítenek a tanulóadatokra visszavetítve, mint a teszteléshez elkülönített adatokon. Ha a rendelkezésünkre álló adatokat rétegzett random módon osztjuk el tanuló és tesztadatokba, akkor feltételezhetjük, hogy az információ aránya az adatpontokhoz képest egyforma. A modell teljesítményének romlása, tehát a valós hibákra vezethető vissza. Minél kisebb a teljesítmény-romlás, annál jobbnak tekinthető a modell.

Ezt a logikát alkalmazva vezettük be az AUC esés (AUC differenc, AUCD) mértéket, amely a széles körben elterjedt ROC görbe alatti terület (Area Under the ROC Curve, Hanley és McNeil 1982) mértéken alapszik. Az AUCD a tanulóadatokon és a tesztadatokon mért AUC különbsége.

A legtöbb jószágérték, így az AUC és ezáltal az AUCD is a modellbecslések által elkövethető kétféle hibát egyformán súlyozza. A PNV-értékelésben ez nem szerencsés, hiszen a „hibás” jelenlétebecslések valódi hibatartalma valószínűleg jóval kisebb, mint a hibás hiánybecsléseké, mégha az utóbbiakban is van információtartalom is. Ezért bevezettünk egy olyan mértéket, amely a hibás jelenlétebecsléseket egyáltalán nem vizsgálja. A Mean Probability Drop (MPD) a megfigyeléseken belül kapott átlagvalószínűség különbsége a tanuló és tesztadatok között.

3. táblázat. A modellbecslések és megfigyelések összehasonlítására szolgáló kontingencia-tábla.

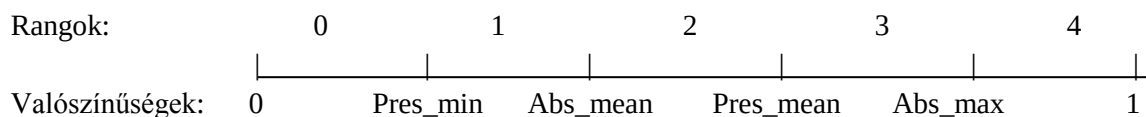
		Becslés	
		1	0
Megfigyelés	1	Eltalált jelenlét (True positive, TP)	Hibás hiánybecslés (False negative, FN)
	0	Hibás jelenlétebecslés (False positive, FP)	Eltalált hiány (True negative, TN)

Egyedi valószínűségi térképek összehasonlíthatóvá tétele

Minden élőhely becsült valószínűségei sajátos eloszlást követnek. Az adott élőhelynél maradva jól értelmezhetők mint alkalmasabb/kevésbé alkalmas területek. Ez azonban élőhelyek között már nem igaz. Esetünkben pl. azonos MÉTA kvadrátban különböző élőhelyek becsült valószínűségei nem jelentenek alkalmassági fokot. Egyes mára megritkult élőhelyeknél (pl. löszölgyes) a becsült valószínűségek még akkor is meglehetősen alacsony abszolút értéket vesznek fel, ha a saját eloszlásához képest már magasak. Ezért egy átskálázási sémát (1. ábra) dolgoztunk ki, amely a megfigyelt jelenlét és hiányadatokon belüli valószínűségek alapvető statisztikáin alapul (4. táblázat). A bevezetett rendszer a legtöbb, átlagos modellre monoton átskálázást biztosít, bár extrém esetek elképzelhetők. A mi esetünkben ilyenek nem fordultak elő.

4. táblázat. A becsült valószínűségek alapvető statisztikái a jelenlét- és hiánymegfigyeléseken belül. Az egyes statisztikákat élőhelyenként számoljuk.

A becsült valószínűségekre vonatkozó statisztikák	Megfigyelések	
	Jelenlét	Hiány
Minimum	Pres_min	
Átlag	Pres_mean	Abs_mean
Maximum		Abs_max

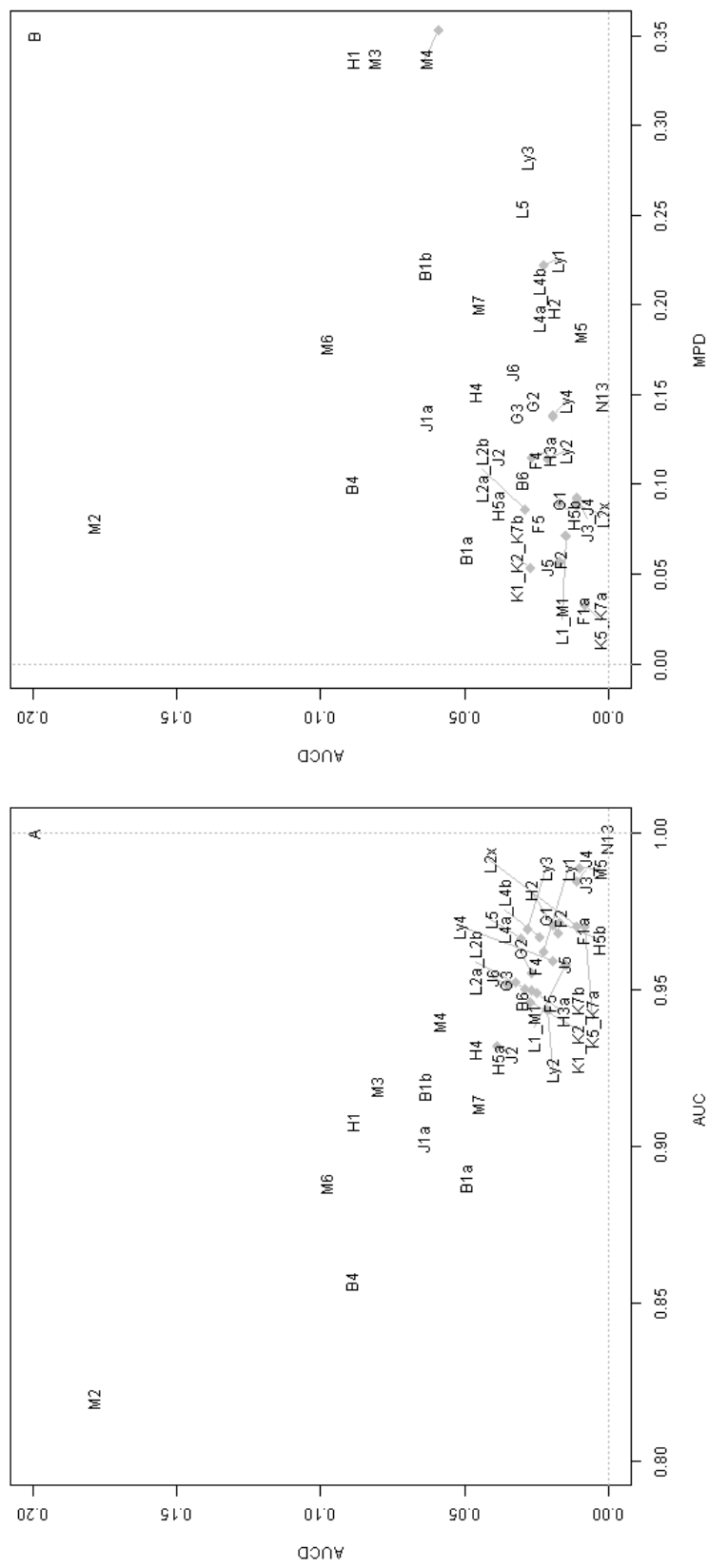


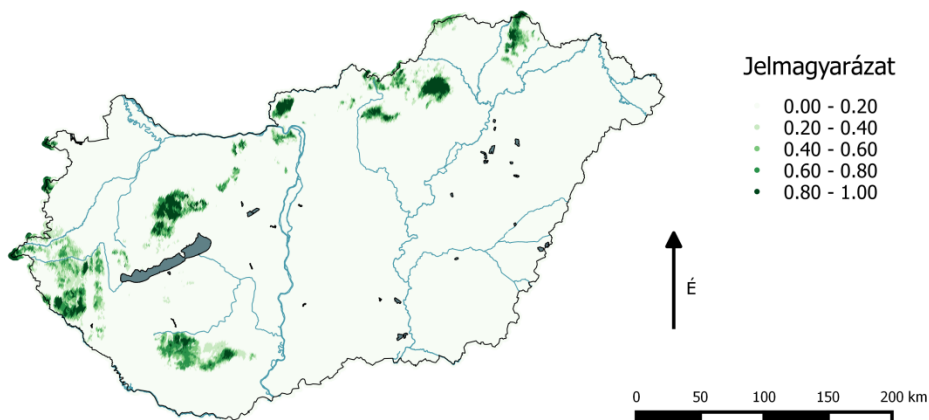
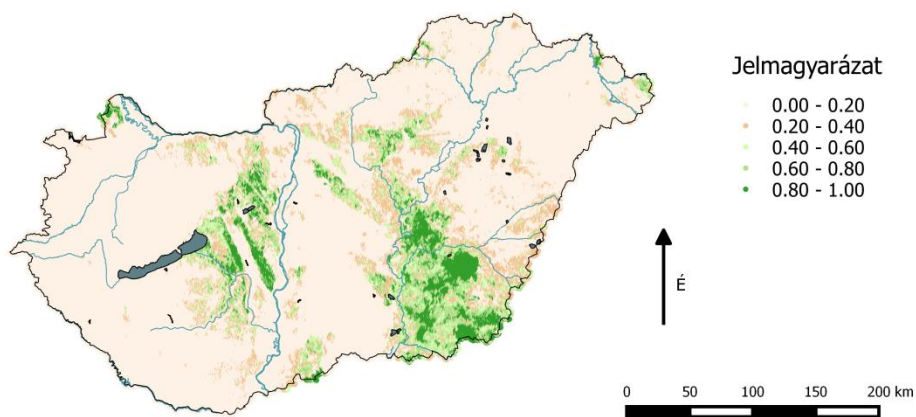
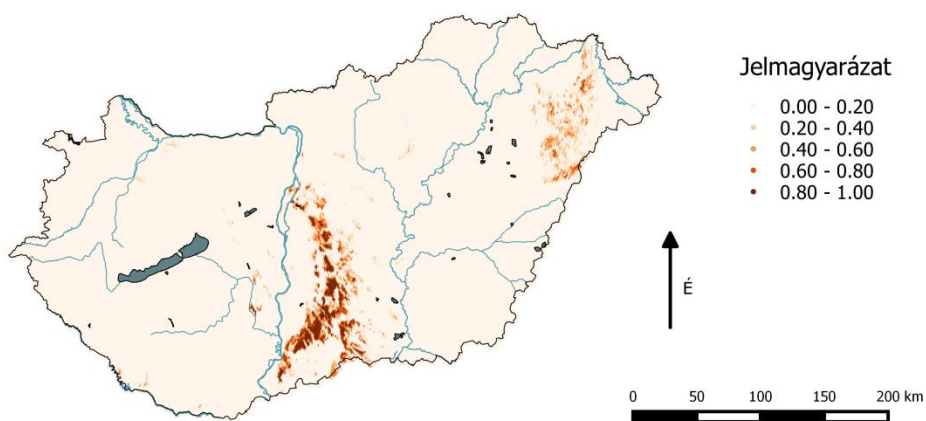
1. ábra. A becsült valószínűségek átskálázási rendszere.

A modellek a hagyományos AUC mérték alapján a kiváló kategóriába sorolhatók (1. táblázat, Swets 1988). A két újonnan bevezetett mérték hasznosnak bizonyult, mert a modelljóságról pusztán az AUC-nél többet árult el. Az AUCD esetén bár erős volt az egybeesés az AUC által adott jóság-sorrenddel, egyes esetekben (pl. B1a vs. M6) megmutatott olyan eltéréseket, amiket a puszta AUC nem. Az MPD más aspektusát mutatta meg a modellek jóságának, például az alföldi löszölgyesek (M2) eszerint meglehetősen megbízható modellt adott. Ebből arra következtethetünk, hogy a rossz AUC alapú eredmények annak tudhatók be, hogy sok a hibás jelenlétebecslés, mert sok helyütt része a potenciális vegetációeloszlásnak. Viszont a hibás hiánybecslések kevésbé jellemzőek. PNV szempontból tehát a lösz erdőssztyepp (M2) modellje is megbízható, ami fontos a hazai PNV-becslés szempontjából, hiszen várható volt, hogy a jelenleginél jóval nagyobb lehetne a kiterjedése.

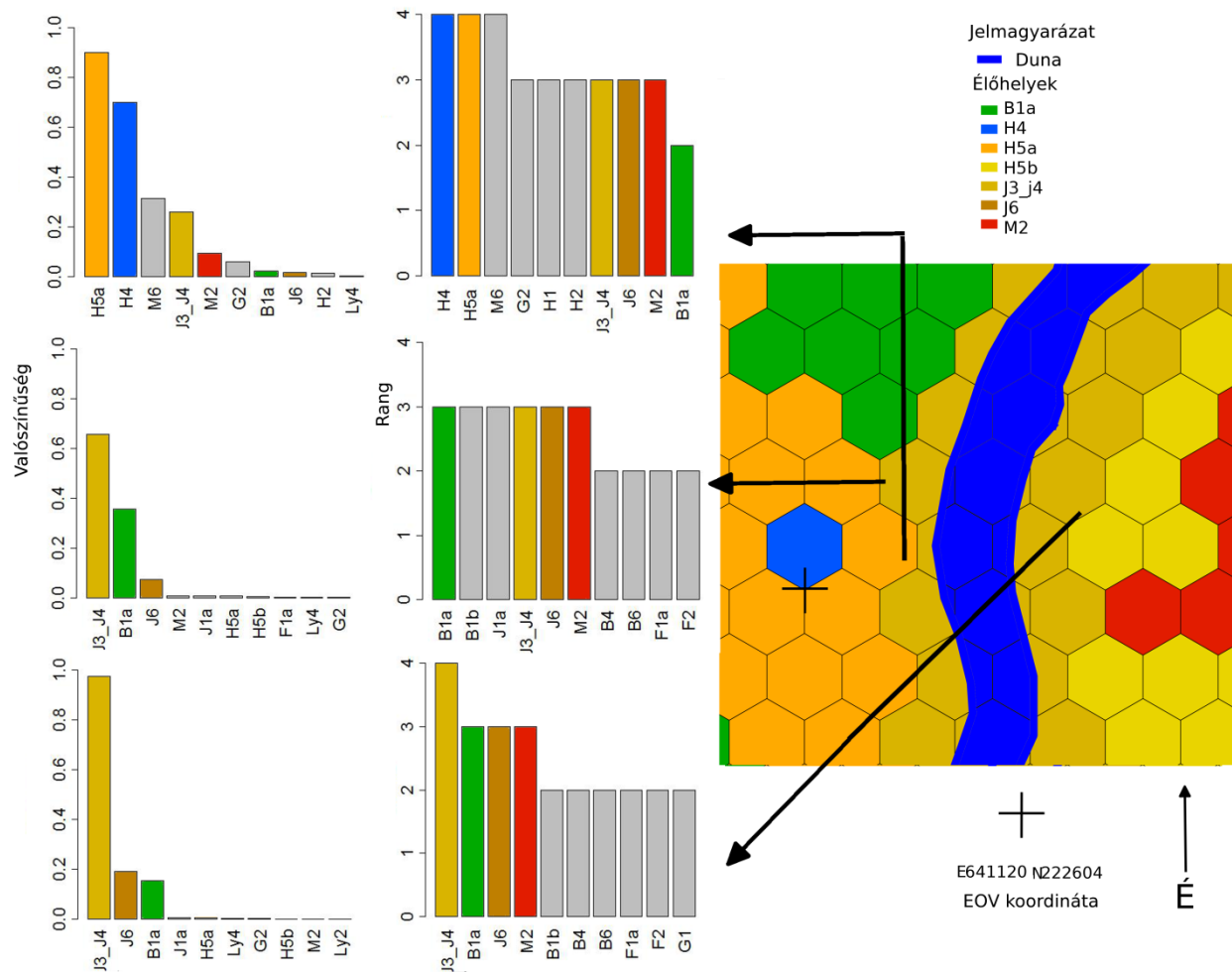
A munka elsődleges eredményei a potenciális elterjedés térképek (3. ábra). Amennyiben egyedi élőhelyek előfordulási valószínűsége a kérdés, ezek a legalkalmasabb becslések, hiszen ezek tartalmazzák a legtöbb információt. Ha azonban egy helyszín, MÉTA hatszög PNV-jére vagyunk kíváncsiak, akkor az átskálázott értékeket kell használnunk. A 4. ábra szemlélteti az MPNV megközelítés előnyeit. A térképi ábrázoláson a hagyományosan alkalmazott determinisztikus PNV látszik, egy gyakori egyszerűsítési megoldással: minden helyen a legnagyobb valószínűséget kapott élőhely feltüntetésével. Látható, hogy ez jelentős információvesztést jelent és hibás következtetésekre is vezethet. Például a löszgyep (H5a) PNV-jű (narancssárga) hatszögek között megjelenik egy erdőssztyepp (H4) PNV-jű hatszög (kék), amely ezáltal eltérőnek tűnik a közvetlen környezetétől. Pusztán a valószínűségeloszlás alapján látható, hogy valójában a közvetlen mellette lévő löszgyep-PNV-jű hatszögben is a következő legvalószínűbb élőhely az erdőssztyepp, vagyis a kétféle hatszög nagyjából hasonló PNV szempontjából, a véletlenül múlik, hogy melyik élőhely kerül az első helyre.

Az ilyen véletleneket és a már fent említett élőhelyspecifikus valószínűségeloszlás-egyenetlenségeket küszöböli ki az átskálázás. Ezután a löszgyep PNV-jű hatszögben immár egyforma, egyben a „legvalószínűbb” kategóriába kerültek a különböző erdőssztyepphez köthető élőhelyek, sőt az alföldi löszölgyes (M2) jelenléte is összemérhető ezekkel, holott a nyers valószínűségek alapján meglehetősen hátra került. Mindez jól összegyeztethető a helyi viszonyokkal, hiszen a kivágat az érdi Fundoklia-völgy oldalát és a Duna két partját mutatja. A dunaparti vegetációra nézve szintén informatívabb az eloszlás-alapú megközelítés, így a legvalószínűbb puhafaligetek (J3_J4) mellett megjelennek a keményfa-ligeterdők is (J6).





3. ábra. Néhány példa egyedi élőhelyek potenciális elterjedési térképeire. a) nyílt homoki gyepek (G1), b) Lössgyepek, kötött talajú sztyeprétek (H5a), c) bükkösösök (K5_K7a) előfordulási valószínűségei.



4. ábra. Kivágat Érd könyékéről, a Funduklia-völgy és a Duna két partja. A térképi kivágat a legvalószínűbb élőhely szerint van színezve. Bal oldalon három egyedi hatszög MPNV-eloszlása látható. Az élőhelyek rövidítéseit az 1. táblázat tartalmazza.

Magyarország MPNV modelleredményeiről egy magyar és egy nemzetközi konferencia-előadást tartottam (Somodi és mtsai. 2015b és Somodi és mtsai. 2014b). A természetvédelmi felhasználási lehetőségekkel (lásd alább) kiegészítve nemzetközi referált folyóiratba kéziratot nyújtottam be (Somodi és mtsai. 2016b).

3. Az MPNV gyakorlati alkalmazása

A PNV becsléseket eddig is széles körben alkalmazták a tájértékelésben és a természetvédelemben. Az MPNV megközelítés azonban egyértelműen szélesebb alkalmazási lehetőségeket biztosít, mint a determinisztikus megközelítés (Somodi és mtsai. 2015a, 2016b). Társszerzőimmal javaslatot tettünk az MPNV alkalmazására a tájértékelésben (Somodi és mtsai. 2015b), a természetvédelemben és az előhely-restaurációban (Somodi és mtsai. 2014a, Török és mtsai. 2015). A kertészeti stilizálás terén a PNV, mint olyan figyelembe vétele is újnak tekinthető (Bede-Fazekas és Somodi 2016).

Természetvédelmi prioritizálás

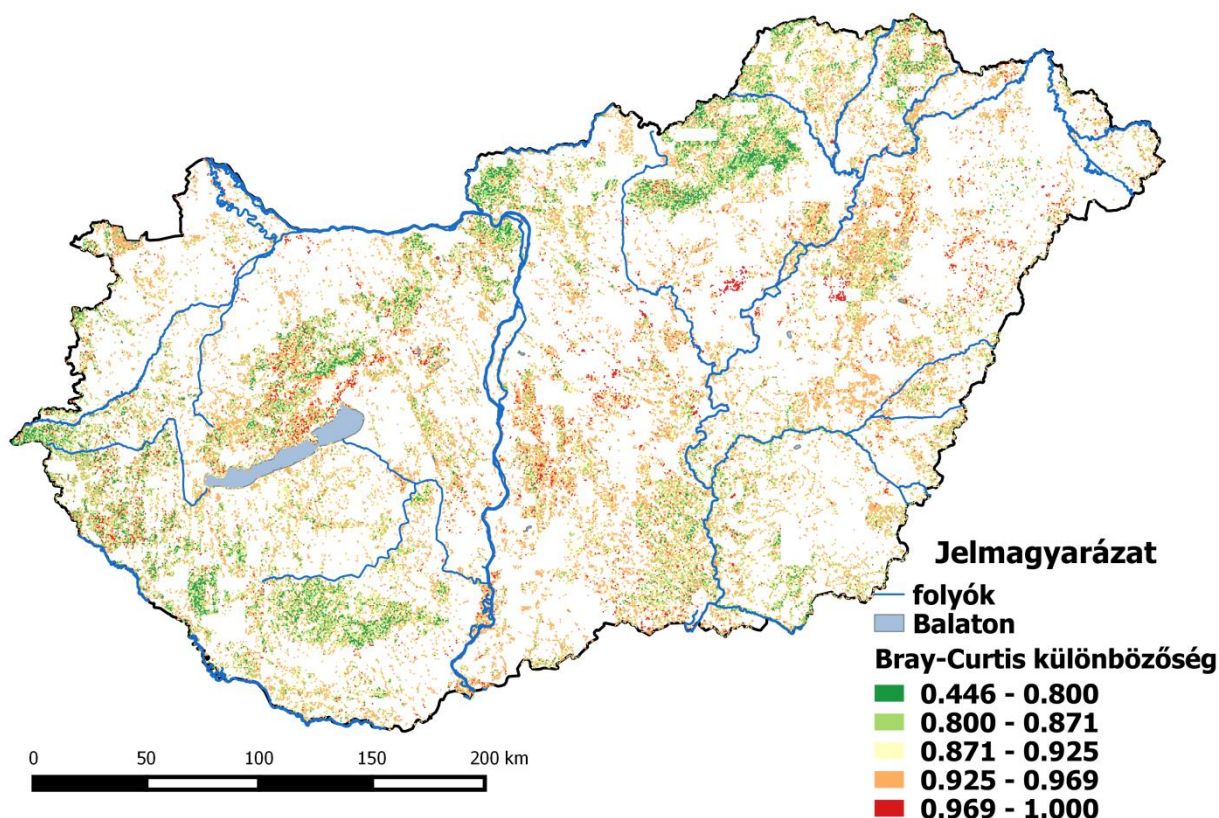
A természetvédelem számára fontos a költséghatékonyság (Brooks és mtsai., 2006, Hoekstra és mtsai., 2005, Reynolds és Hessburg, 2005, Wilson és mtsai., 2007), hogy a társadalom által a célra szánt keretből a lehető legtöbb élőhelyet meg lehessen őrizni. A védelemben a PNV alapú prioritássorrend kialakítása költséghatékony megoldás, mert a PNV éppen azt fejezi ki, hogy az adott terület aktuális környezeti adottságai megfelelnek az adott élőhelynek. Amennyiben ettől eltér a védelem tárgya (amire lehetnek érvényes indokok), akkor vállalni kell a folyamatos, emberi beavatkozást, melynek költségigényével is számolni kell. Így van ez akkor is, ha csak a prehumán állapotokat tartja szem előtt, hiszen az akkori környezet eltérhet és gyakran el is tér a maitól.

Az MPNV megközelítés előnye, hogy nagyobb szeletét látni engedi a környezet által fenntartható vegetációtípusoknak. Előfordulhat ugyanis, hogy a megvalósult élőhely csak egy a lehetséges PNV-k közül, ami esetleg elsikkadna a determinisztikus becslésnél, ahogy ezt a 4. ábra is szemléltette.

Restauráció

Itt is a fenntarthatóság és költséghatékonyság szempontja indokolja a PNV-alapú célállapot-kiválasztást. Az EU 2020 Biodiverzitás Stratégiában is szereplő Aichi 15 célkitűzés (azaz, hogy a jelenlegi degradált élőhelyek 15%-át restaurálják; CBD, 2011), csak úgy közelíthető meg, ha olyan élőhelyeket hoznak létre, amelyek a kezdeti kialakítástól eltekintve a minimális emberi beavatkozást igénylik. Ennek a legjobb közelítése a PNV-becsítés. A restaurációs tervezés számára fontos szempont a megvalósíthatóság (Bakker és Berendse, 1999, Ehrenfeld, 2000) és a társadalmi elfogadottság (Buckley és Crone, 2008, Pfadenhauer, 2001). Amennyiben egy determinisztikus PNV becsítés áll csak rendelkezésre, a kivitelező elvetheti a természetes állapot megcélzását, amennyiben az adott vegetációtípus nehezen restaurálható. Az MPNV becsítés előnye, hogy több lehetséges céltípust megmutat, amelyekhez kategóriákat rendel. Így az optimális tervezésnél van természetes választási lehetőség.

Tájaink természetességének felmérése gyakori kérdés (Ferrari és mtsai., 2008, Renetzeder és mtsai., 2010, Rüdisser és mtsai., 2012). Többek között mivel a természetvédelem kezd elmozdulni egyedi állományok védelmétől az élőhelykomplexek védelme irányába (Angelstam és mtsai., 2004, Lindenmayer és Fischer, 2013, Opdam és mtsai., 2002). Cél lehet az ezeket leginkább reprezentáló tájrészletek védelme, vagy a hiányzó élőhelyek restaurációja a kevésbé természetes tájrészletekben. Itt az MPNV tud összehasonlítási alapot nyújtani, hogy egy-egy táj, tájrészlet mennyire valósítja meg a környezeti háttér adta lehetőségeket (5. ábra). Az MPNV használata ilyen esetben jobban szolgálja a fenntarthatóságot, mint a ma még elterjedtebb megközelítés: a prehumán viszonyok alapul vétele, hiszen Magyarországon, de Európában másutt is az ember tájátalakító tevékenysége óta az élettelen környezet is megváltozott, részben természetes folyamatok (történeti klímaváltozás), részben az emberi tevékenység (folyószabályozás, talajvízcsökkenés) hatására.



5. ábra. A megfigyelt aktuális természetes élőhely-eloszlás és a becsült MPNV közötti eltérés.

A kertészeti értelemben vett stilizálás során dísznövényekkel (amelyek között lehet őshonos) idéznek meg természetes élőhelyeket városi környezetben kertekben, parkokban. Célja eltér az élőhely-restaurációtól, márcsak azért is mert olyan kis helyeken igyekszik az élőhely imitálására, ahol önfenntartó állomány létrehozása lehetetlen volna. Nem is törekszik tehát erre. Elsődleges szempont a képi hasonlóság, nem pedig a fajkészlet összetétele, különösen ha az eredeti élőhely környezeti igényei nagyon eltérnek a megvalósítás helyétől. Ilyenkor az eredeti élőhely helyben is alkalmas fajai mellett az eredeti fajkészletet megidéző dísznövények telepítése történik. Leggyakrabban mediterrán kerteket hoznak létre ilyen módon a mediterránumon kívül, a hazai élőhelyek közül eddig nyáras-borókások és bokorerdők megidézésére merült fel eddig (Schmidt, 2003).

Felvetettük, hogy a stilizálás alkalmas lehet klímaváltozás miatt eltűnő élőhelytípusok megidézésére (Bede-Fazekas és Somodi, 2013, 2016). A PNV modellek változószerkezete alapján valószínűsítettük a fenyőelegyes erdők klímaérzékenységét, ezért ennek a példáján mutattunk be az emlékörzés jellegű stilizálási lehetőséget. Mivel fás élőhely, a karakter kialakításában a fásszárúak jellege meghatározó. Azonosítottuk a fafajokat, amelyek a változó klíma mellett is megtarthatók, míg az érzékenyebbek helyett hasonló kinézetű, de szárazságtűrőbb fajokat javasoltunk, mivel a hazai klímaváltozási forgatókönyvek szerint a tényszerű időszakban szárazodás várható (Bartholy et al. 2007; Bartholy és Pongrácz 2008). Az állományalkotó fajok közül az erdeifenyő (*Pinus sylvestris*) és a pionír lombosfajok (*Betula pendula* és *Populus tremula*) a klímaváltozás fényében is telepíthetők, míg pl. a közönséges luc (*Picea abies*) és a bükköt (*Fagus sylvatica*) feltehetően helyettesíteni kell. A luc megidézésére a koreai luc (*Picea koyamae*) illetve egyes jegenyefenyő-fajok (*Abies numidica*, *Abies cilicica*) lehetnek alkalmasak. Mivel a keleti bükk (*Fagus orientalis*) elterjedése alapján a büknél szárazságtűrőbb, alkalmas lehet az utóbbi megidézésére. Az aljnövényzet kialakítása a természetes állományok fajainak talajigénye biztosításával és a faszervezet létrehozta megfelelő árnyékolás, ill. öntözés kombinációjával akár az eredeti fajok megtartásával is megoldható.

4. A modellértékelés egy aspektusa: a True Skill Statisztika (True Skill Statistics, TSS) kritikája

A modellértékelési kihívásokkal való megküzdés első logikus lépése az irodalom áttekintése az aktuálisan elérhető módszerek felkutatása, tesztelése volt. Az utóbbi során derült fény arra, hogy az egyik, ma kifejezetten elterjedt (pl. Coetzee és mtsai. 2009, Jones és mtsai. 2010; Zurell és mtsai. 2012, Engler és mtsai. 2013, Pliscoff és mtsai. 2014) mérőszám, a TSS nem rendelkezik azzal a tulajdonsággal, amit az irodalomban neki tulajdonítanak. Mélyebben megvizsgálva a kérdést kiderült, hogy a TSS-t az ökológiában a korábban elterjedt Kappa (Cohen 1960) alternatívájaként bevezető szerzők (Allouche et al. 2006) szimulációs stratégiája vezetett hibás következtetésre.

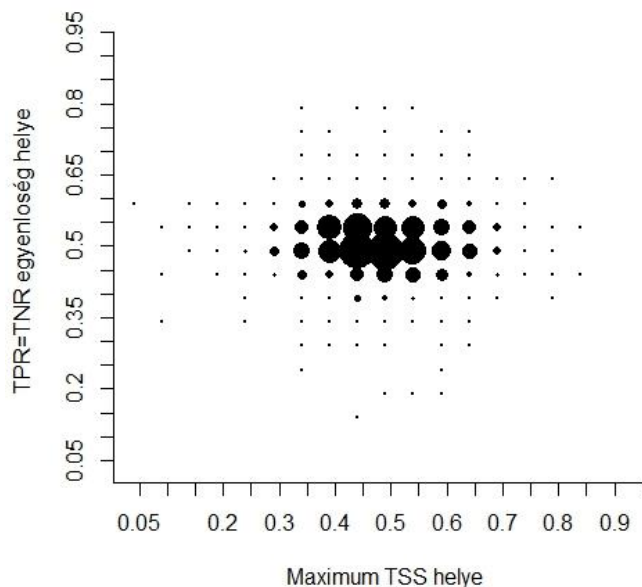
Allouche és mtsai. (2006) azt állították, hogy a TSS értéke csak a modell jóságát tükrözi (ahogy az kívánatos) és nem függ a prevalenciától, azaz a bementi adatokban található jelenlét ill. hiánymegfigyelések arányától. Az általuk tapasztalt függetlenség miatt ajánlották a TSS-t a Kappával szemben, amelynek prevalencia-függése már korábban ismert volt. Ez vezetett a TSS elterjedéséhez az ökológiában. Mind elméleti, mind szimulációs ellenérveket találtunk.

Elméleti szempontból úgy találtuk, hogy Allouche és mtsai. (2006) csak egy speciális esetet vizsgáltak, amikor is a TSS maximuma mindig ott van, ahol két másik modellértékelésben szokásos érték, a TPR és TNR ($TPR = TP / (TP + FP)$, $TNR = TN / (TN + FN)$, 3. táblázat) éppen egyenlő. Más esetre Allouche és mtsai. (2006) megközelítése nem releváns. Szimulációinkban bebizonyosodott, hogy ez az egybeesés nem feltétlen valósul meg (6. ábra).

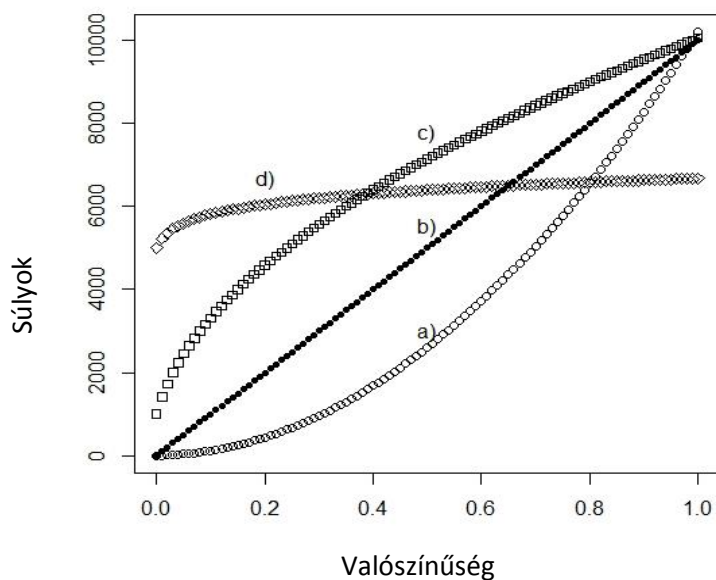
Szimulációink során mesterséges becsléseket állítottunk elő, úgy hogy meghatározott prevalencia-szinteknél azonos függvények mentén állítottunk elő becsléseket, oly módon, hogy a megfigyeléshez beállított megfigyelt jelenlétekhez, ill. hiányokhoz a függvények alapján kiosztott súlyok szerint rendeltünk becsült valószínűségeket (7. ábra). A függvények voltak hivatottak a modelljóság modellezésére.

100 futás eredményeit átlagolva, azt tapasztaltuk, hogy a TSS értéke változik a prevalencia változásával. A változás mértéke a prevalencia extrém értékei (alacsony, magas) és a gyengébb modellek felé nő (8. ábra).

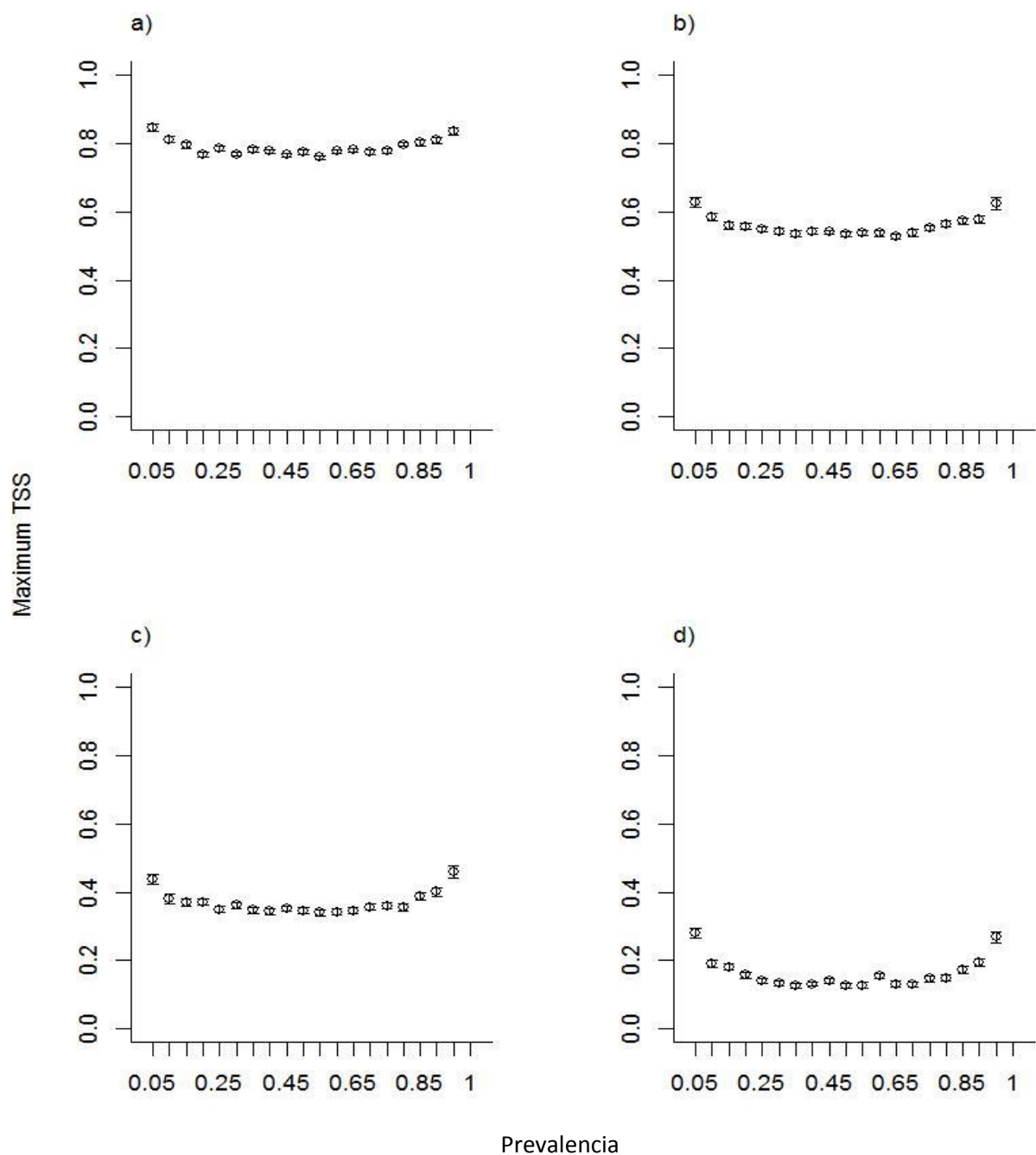
A TSS prevalenciafüggésének tisztázásra készült munkánk bírálat alatt van a *Methods in Ecology and Evolution* folyóiratban (Somodi és mtsai. 2016a).



6. ábra. A TSS maximumának helye és a $TNR=TPR$ egyenlőséghez tartozó vágási pont nem feltétlenül esik egybe. A körök mérete az esetek gyakoriságával arányos. 100 ismétlés eredménye. A négyzetes valószínűségkiosztást használó szimuláció ezirányú eredményei alapján mutatjuk be, a más kiosztással dolgozó szimulációk hasonló eredményt adtak.



7. ábra. A becsült valószínűségekhez kiosztásához használt függvények alakja. Itt a súlyok emelkedése a szimulációban szereplő jelenlétekre vonatkozik. a) Négyzetes, b) lineáris, c) gyök, d) -16. hatvány függvény. Minél magasabbak az értékek a 0-hoz közeli tartományban, annál gyengébb a modell.



8. ábra. A TSS maximumának változás a prevalencia (x tengely) változásával. A részabrák különböző modelljóságoknak felelnek meg, az ennek szimulációjához használt függvények: a) négyzetes, b) lineáris, c) gyök, d) -16. hatvány függvény. 100 futás eredménye, az ábrán az átlag és az átlag szórása látható.

5. A mintázatelemzés előfeltétele: valószínűségbecslések binarizálásának lehetőségei

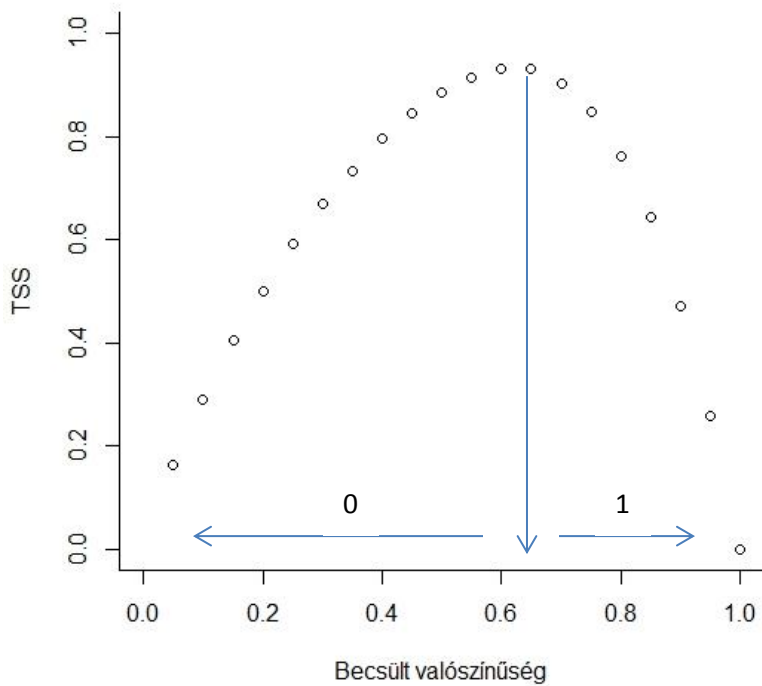
A posztdoktori kutatás utolsó tervezett pontja a becsült és megfigyelt mintázatok összehasonlító elemzése volt. A valószínűségi becslések binarizálása természetesen információvesztéssel jár, ezért ezt ahol lehet elkerülendőnek tartom. Ezért is dolgoztuk ki az MPNV keretrendszert, illetve az élőhelyek valószínűségének összefésülési lehetőségét. Ugyanakkor a térbeli mintázatok elemzésének standard módszerei jelenlét-hiány (vagy kategoriális, de még csak nem is ordinális) mintázatokon alapulnak. A szakirodalmi összehasonlításhoz feltétlen szükséges, hogy akár új módszerek mellett is az elterjedt módon is megtörténjen a mintázatelemzés, így szükség van a becslések binarizálására.

A binarizálás témakörében rendkívül szerteágazó a szakirodalom, a benne való elmélyedéssel kiderült, hogy nem állpítható meg egyszerűen egy legalkalmasabb módszer. Az átfedések és szinonímák nagy száma jelentősen nehezíti az eligazodást. Emiatt a konkrét mintázatelemzést megelőző feladatként körvonalazódott a binarizálási lehetőségek áttekintése, a szinonímák feloldása és az azonos eredményre vezető indexek azonosítása. Az eredményekről nemzetközi referált folyóiratban tervezünk reviewcikket megjelentetni (Lepesi és mtsai. 2016). A kézirat a szöveg szerkesztésének fázisában van.

A binarizálás alapelve, hogy a becsült valószínűségek skáláján felfelé haladva megkeressük azt az optimális vágási pontot, ahol a leghatékonyabban bontható a valószínűségi gradiens becsült jelenlétekre és hiányokra (9. ábra). Ez többnyire a már említett kontingencia-táblán (3. táblázat) alapuló módszerrel számítják, oly módon, hogy a választott vágási pontnál legyen maximális a becslés siker.

Binarizálási módszerek számos más tudományágban is előfordulnak, mi áttekintésünket az adott ökológiai problémákra alkalmazott megoldásokra szűkítettük a kezelhetőség érdekében. Az áttekintés során egyértelművé vált az átfedések nagy száma, ezért külön hangsúlyt fektettünk a szinonimák azonosítására. Az indexek leírása szintaktikailag is meglehetősen változatos volt, ezért egységesítettük a képleteket a faj-előfordulás modellekhez általánosan használt kontingencia-tábla (3. táblázat) konvencióit használva. Ezáltal fény derült a képletből adódóan azonos vágáspontot eredményező megközelítésekre is.

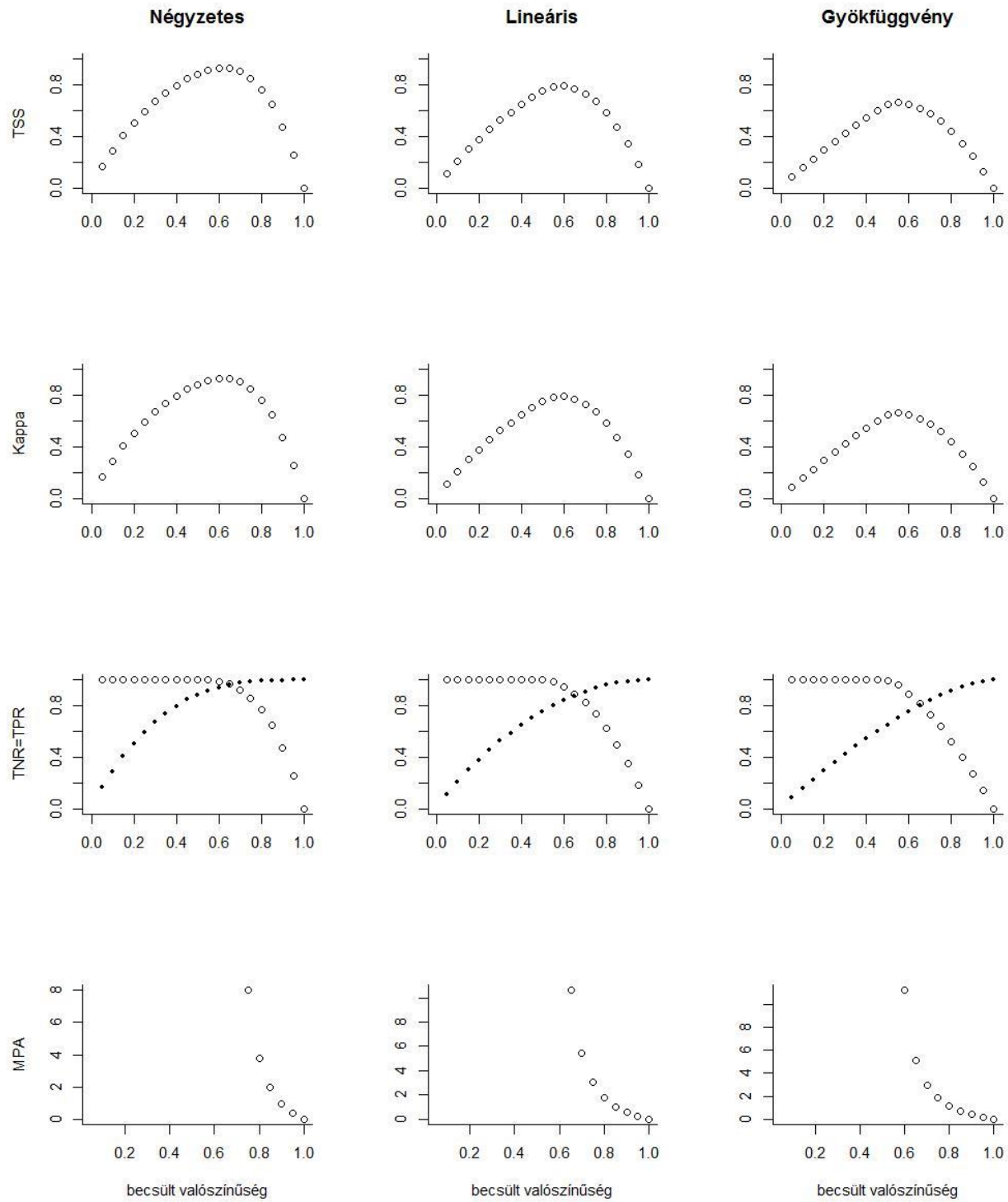
Ennek eredményeképpen 4 alapesetre tudtuk visszavezetni az irodalomban előforduló binarizációs megoldásokat (5. táblázat). Ennek a 4 indexnek a vágópontoktól függő viselkedését szimulációval is teszteltük. A szimuláció elve a TSS-tesztelésnél leírt volt, azaz itt is a 7. ábrán bemutatott függvények (kivéve -16. hatvány) lefutásának megfelelően allokáltunk becsült valószínűségeket a hiány és jelenlét-„adatok”-hoz. Ebben a vizsgálatban azonban a prevalencia-függést nem vizsgáltuk, mert a vágási hely változása a prevalenciával szükségszerű. Ismét 100 ismétést alkalmaztunk (10. ábra).



9. ábra. A binarizálás szemléltetése a TSS példáján. A binarizáláshoz használt mértéket a lehetséges vágópontok mentén kiszámoljuk és a valószínűségeket ott osztjuk becsült hiányokra és jelenlétekre, ahol a mérték felveszi a meghatározott értéket. Ez a TSS esetében a maximumhely, ezt nyíl jelöli az ábrán.

5. táblázat. Az irodalomban fellelhető binarizálási módszerek. A módszerek visszavezethetők 4 alapesetre, melynek szinonimáit és kontingencia-táblázaton alapuló képletét is feltüntettük. Emellett megadtuk azokat a módszereket, amelyek bár már képlettel, de ugyanarra a vágási helyre jutnak.

Név	Szinonima	Képlet	Azonos vágási helyre vezető módszerek
True Positive Rate (TPR) = True Negative Rate (TNR) (Liu és mtsai. 2005)	sensitivity/specificity (Fielding és Bell 1997) equality point	$TPR = TP / (TP + FN)$, $TNR = TN / (FP + TN)$	MDT (minimized difference threshold) MDT = TPR - TNR (Rojas és mtsai. 2001)
Kappa (Cohen 1960) maximuma	Overall prediction success (Manel és mtsai. 2001), Heidke skill score (Stephenson 1999), Módosított Rand-index (Saltstone és Stange 1996)	$\frac{((TP + TN)/n) - (TP + FP)(TP + FN) + (FN + TN)(TN + FP)}{n^2}$ $(1 - \frac{((TP + FP)(TP + FN) + (FN + TN)(TN + FP))}{n^2}))$	
True Skill Statistic (TSS) Allouche és mtsai. (2006) maximuma	Youden index (Youden 1950)	$TSS = (TP / (TP + FN)) + (TN / (FP + TN)) - 1$	Maximized Sum Threshold (MST) MST = TPR + TNR (Manel és mtsai. 2001)
Minimal Predicted Area (MPA) (Engler és mtsai. 2004)		$TP > (TP + FP) * 0,9$	



10. ábra. A főbb binarizációs mértéktípusok viselkedése a szimulációkban. 100 futás átlagai. Az oszlopokban a modellbecslések kiosztásához használt különféle függvények szerint csoportosítottam az ábrákat. A modelljószág jobbról balra csökken.

7. Köszönetnyilvánítás

Az OTKA PD 83522 projekt mellett köszönettel tartozom Magyarország Élőhelyeinek Térképi Adatbázisa (MÉTA)-nak, hogy a modellépítéshez elérhetővé tették az adatbázisban szereplő adatokat. Valamint a témában készült cikkek társszerzőinek.

8. Irodalomjegyzék

A pályázat támogatásával elkészült publikációk (az OTKA mindegyikben fel van tüntetve támogatóként)

Nemzetközi folyóiratcikkek:

Somodi I., Molnár Zs., & Ewald J. (2012) **Towards a more transparent use of the Potential Natural Vegetation concept – an answer to Chiarucci et al.** Journal of Vegetation Science, 23, 590-595.

Fekete G., Molnár Zs., Magyar E., Somodi I., & Varga Z. (2014) **A new framework for understanding Pannonian vegetation patterns: regularities, deviations and uniqueness.** Community Ecology, 15, 12-26.

Bede-Fazekas Á., & Somodi I. (2015) **Stylization: a method for preserving the character of climate sensitive habitats. A case study of acidofrequent mixed forests.** Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis, 63, (közlésre elfogadva).

Hazai folyóiratcikkek:

Fekete G., Molnár Zs., Magyar E., Somodi I., & Varga Z. (2011) **Egyediség, szabályosság és deviáció a Pannon régió vegetációjának példáján.** Botanikai Közlemények, 98, 29-59.

Bede-Fazekas Á., & Somodi I. (2013) **Mészkerülő lombelegyes fenyvesek mint éghajlatérzékeny élőhelyek kertépítészeti stilizálása.** 4D Tájépítészeti és Kertművészeti Folyóirat, 8, 54-71.

Nemzetközi konferencia-absztraktok:

Somodi I., Czúcz B., Molnár Z., & Zimmermann N.E. (2014a). **Multiple Potential Natural Vegetation (PNV) assessment – an introduction through the modelling of the PNV of Hungary.** In: Čarni A, Juvan N, Ribeiro D (szerk.) 23rd International Workshop of the European Vegetation Survey: Book of Abstract. p. 55

Somodi I., Zimmermann N.E., Molnár Zs., Czucz B., & Ewald J. (2014b) **Multiple Potential Natural Vegetation modelling an assistance for regional scale planning of restoration.** In: Anne Tolvanen, Anne-Maarit Hekkala (szerk.) 9th European Conference on Ecological Restoration, p. 134.

Somodi I., Lepesi N., & Botta-Dukát Z. (2015a) **Warning of the prevalence dependence of the True Skill Statistics – a frequently used model goodness measure in vegetation modelling.** In: Milan Chytrý, David Zelený, Eva Hettenbergerová (szerk.) 58th Annual Symposium of the International Association for Vegetation Science, p. 349

Hazai konferencia-absztraktok:

Somodi I., Bölöni J., & Molnár Zs. (2015b). **Magyarország többretegű potenciális vegetációbecslésének (PNV) lehetséges alkalmazásai.** In: Padisák Judit, Liker András, Stenger-Kovács Csilla (szerk.) X. Magyar Ökológus Kongresszus. p. 126.

Török K., Csecserits A., Halassy M., Kövendi-Jakó A., Somodi I., & Dezsényi P. (2015) **Természetközeli élőhelyek létrehozása iparterületen.** In: Padisák Judit, Liker András, Stenger-Kovács Csilla (szerk.) X. Magyar Ökológus Kongresszus. p. 153

Nemzetközi folyóiratcikkek előkészületben:

Somodi I., Lepesi N., & Botta-Dukát Z. (2016a)*. **True Skill Statistics is not free of prevalence bias – a critique of Allouche et al. (2006).** Methods in Ecology and Evolution (bírálat alatt)

Somodi I., Molnár Zs., Czucz B., Bede-Fazekas Á., Bölöni J., & Zimmermann N.E. (2016b) **Implementation, evaluation and application of Multiple Potential Natural Vegetation models from a conservation perspective.** (benyújtva)

Lepesi N., Botta-Dukát Z., & Somodi I. (2016) **Determining an optimal binarisation threshold in predictive ecological modelling.** (előkészületben)

* A valószínűsíthető évszámot a szövegbeli hivatkozás egyértelműsítése végett adtam meg.

A szövegben idézett egyéb munkák:

Allouche, O., Tsoar, A., & Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). Journal of Applied Ecology, 43, 1223-1232.

- Angelstam, P., Roberge, J.-M., Lõhmus, A., Bergmanis, M., Brazaitis, G., Dönn-Breuss, M., Edenius, L., Kosinski, Z., Kurlavicius, P., Lärmanis, V., Lūkins, M., Mikusinski, G., Račinskis, E., Strazds, M., & Tryjanowski, P. (2004) Habitat modelling as a tool for landscape-scale conservation: a review of parameters for focal forest birds. *Ecological Bulletins*, 51, 427–453.
- Attorre, F., Francesconi, F., De Sanctis, M., Alfò, M., Martella, F., Valenti, R., & Vitale, M. (2014): Classifying and mapping potential distribution of forest types using a finite mixture model. *Folia Geobotanica*, 49, 313–335.
- Bakker, J.P. & Berendse, F. (1999) Constraints in the restoration of ecological diversity in grassland and heathland communities. *Trends in Ecology & Evolution*, 14, 63–68.
- Bartholy J., & Pongrácz R. (2008) Regionális éghajlatváltozás elemzése a Kárpát-medence térségére. In: Harnos Zs., & Csete L. (szerk.): *Klímaváltozás: környezet – kockázat – társadalom*. Szaktudás Kiadó Ház, Budapest.
- Bartholy J., Pongrácz R. & Gelybó Gy. (2007) A 21. század végén várható éghajlatváltozás Magyarországon. *Földrajzi Értesítő*, 56, 147–168.
- Brooks, T.M., Mittermeier, R.A., Fonseca, G.A.B., Gerlach, J., Hoffmann, M., Lamoreux, J.F., Mittermeier, C.G., Pilgrim, J.D. & Rodrigues, A.S.L. (2006) Global biodiversity conservation priorities. *Science*, 313, 58–61.
- Buckley, M.C. & Crone, E.E. (2008) Negative off-site impacts of ecological restoration: understanding and addressing the conflict. *Conservation Biology*, 22, 1118–1124.
- CBD (2011): Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity. Secretariat of the Convention on Biological Diversity. Montreal, Canada. pp. 25.
- Chiarucci, A., Araújo, M.B., Decocq, G., Beierkuhnlein, C., & Fernández-Palacios, J.M. (2010): The concept of potential natural vegetation: an epitaph? *Journal of Vegetation Science*, 21, 1172–1178.
- Coetzee, B.W., Robertson, M.P., Erasmus, B.F., Van Rensburg, B.J., & Thuiller, W. (2009) Ensemble models predict Important Bird Areas in southern Africa will become less effective for conserving endemic birds under climate change. *Global Ecology and Biogeography*, 18, 701–710.
- Ehrenfeld, J.G. (2000) Defining the limits of restoration: the need for realistic goals. *Restoration Ecology*, 8, 2–9.
- Engler, R., Guisan, A., & Rechsteiner, L. (2004). An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, 41, 263–274.
- Engler R., Waser L., Zimmerman N.E., Schaub M., Berdos S., Ginzler C. & Psomas A. (2013) Combining ensemble modeling and remote sensing for mapping individual tree species at high spatial resolution. *Forest Ecology and Management*, 310, 64–73.

- Ferrari, C., Pezzi, G., Diani, L., & Corazza, M. (2008): Evaluating landscape quality with vegetation naturalness maps: an index and some inferences. *Applied Vegetation Science*, 11, 243–250.
- Fielding, A.H., & Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental conservation*, 24, 38-49.
- Hanley, J.A., & McNeil, B.J. (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143, 29–36.
- Hemsing, L.Ø. & Bryn, A. (2012) Three methods for modelling potential natural vegetation (PNV) compared: A methodological case study from south-central Norway. *Norsk Geografisk Tidsskrift – Norwegian Journal of Geography*, 66, 11–29.
- Hoekstra, J.M., Boucher, T.M., Ricketts, T.H. & Roberts, C. (2005) Confronting a biome crisis: Global disparities of habitat loss and protection. *Ecology Letters*, 8, 23–29.
- Jones, C.C., Acker, S.A. & Halpern, C.B. (2010) Combining local-and large-scale models to predict the distributions of invasive plant species. *Ecological Applications*, 20, 311-326.
- Lindenmayer, D.B., & Fischer, J. (2013) *Habitat fragmentation and landscape change: an ecological and conservation synthesis*. Island Press: Washington D.C., USA. pp. 328.
- Liu, C., Berry, P.M., Dawson, T.P., & Pearson, R. G. 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28, 385-393
- Manel, S., Williams, H.C. & Ormerod, S.J. (2001) Evaluating presence–absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology* 38, 921–931.
- Opdam, P., Foppen, R., & Vos, C. (2002) Bridging the gap between ecology and spatial planning in landscape ecology. *Landscape Ecology*, 16, 767–779.
- Pfadenhauer, J. (2001) Some remarks on the socio-cultural background of restoration ecology. *Restoration Ecology*, 9, 220–229.
- Pliscoff, P., Luebert, F., Hilger, H.H., & Guisan, A. (2014) Effects of alternative sets of climatic predictors on species distribution models and associated estimates of extinction risk: A test with plants in an arid environment. *Ecological Modelling*, 288, 166-177.
- Reger, B., Häring, T., & Ewald, J. (2014) The TRM-model of potential natural vegetation in mountain forests. *Folia Geobotanica*, 49, 337–359.
- Reynolds, K.M. & Hessburg, P.F. (2005) Decision support for integrated landscape evaluation and restoration planning. *Forest Ecology and Management*, 207, 263–278.
- Renetzeder, C., Schindler, S., Peterseil, J., Prinz, M. A., Mücher, S. & Wrba, T. (2010): Can we measure ecological sustainability? Landscape pattern as an indicator for naturalness and land use intensity at regional, national and European level. *Ecological Indicators*, 10, 39–48.

Rojas, A. B., Cotilla, I., Real, R., & Palomo, L. J. (2001) Determinación de las áreas probables de distribución de los mamíferos terrestres en la provincia de Málaga. *Galemys: Boletín SECEM*, 13, 217-229

Rüdisser, J., Tasser, E. & Tappeiner, U. (2012) Distance to nature – A new biodiversity relevant environmental indicator set at the landscape level. *Ecological Indicators*, 15, 208–216.

Saltstone, R., & Stange, K. (1996) A Computer Program to Calculate Hubert and Arabie's Adjusted Rand Index. *Journal of Classification*, 13, 169-172.

Schmidt G. (2003) Stilizált növénytársulások a kertben. In: Schmidt G. (szerk.): *Növények a kertépítészetben*. Mezőgazda Kiadó, Budapest.

Stephenson, D. B. (2000). Use of the “odds ratio” for diagnosing forecast skill. *Weather and Forecasting*, 15, 221-232.

Tüxen, R. (1956) Die heutige potentielle natürliche Vegetation als Gegenstand der Vegetationskartierung. *Angewandte Pflanzensociologie (Stolzenau)*, 13, 5–42.

Wilson, K.A., Underwood, E.C., Morrison, S.A., Klausmeyer, K.R., Murdoch, W.W., Reyers, B., Wardell-Johnson, G., Marquet, P.A., Rundel, P.W., McBride, M.F., Pressey, R.L., Bode, M., Hoekstra, J.M., Andelman, S., Looker, M., Rondinini, C., Kareiva, P., Shaw, M.R. & Possingham, H.P. (2007) Conserving Biodiversity Efficiently: What to Do, Where, and When. *PLoS Biology*, 5, e223.

Youden, W.J. (1950) Index for rating diagnostic tests. *Cancer*, 3, 32–35.

Zurell, D., Grimm, V., Rossmann, E., Zbinden, N., Zimmermann, N.E., & Schröder B. (2012) Uncertainty in predictions of range dynamics: Black grouse climbing the Swiss Alps. *Ecography*, 35, 590-603.