

Válasz
Dr. Tóthmérész Béla, az MTA Doktora
opponensi véleményére

Először is szeretném megköszönni Bírálómnak, hogy vállalta értekezésem bírálatának elkészítését. Köszönöm a dolgozatról írt dicsérő szavait és külön köszönöm, hogy bírálatában tágabb tudománytörténeti perspektívába helyezi munkámat.

Opponensem a bírálat bevezető részében utal a numerikus ökológia 70-es – 80-as évekbeli hőskorára, amikor a számítógépek elterjedésével sorra születtek a vegetációtudományban alkalmazható sokváltozós elemző módszerek. Pályám kezdetén, még szakdolgozó egyetemistaként magam is sokat olvastam ennek a kornak a cikkeit. A cikkekből sugárzó optimizmus, amely szerint az új módszerek alkalmazása megoldja a vegetációtan korábbi problémáit, rám is hatott, amikor kutatási témát választottam. Azonban én is úgy éreztem, hogy a módszerek már rendelkezésre állnak, és inkább az alkalmazásuk lesz mi generációnk feladata, ami a technika fejlődésének köszönhetően egyre nagyobb adatmennyiségeken történhet. Doktorandusz hallgatóként Fekete Gábor hatására kezdtem el érdeklődni a klasszikus cönológia elméleti alapozása iránt, és ekkor vettem észre, hogy a fidelitás koncepció, amely a klasszikus megközelítés központi eleme, nem került át a numerikus módszerek közé, pedig objektív mérőszámokat alkalmazva ott is hasznos lenne. Azt, hogy ez a megérzésem helyes volt, jól bizonyítja, hogy az ebben a témában cseh kollégákkal közösen megjelent 2002-es cikkünkre 100-nál több független hivatkozást kaptunk. Ahogy Opponensem is látja, számos további módszerfejlesztést a nagy adatbázisok „kényszerítenek ki”, amelyek elemzésekor már nincs lehetőség az eredmények minden részletre kiterjedő értelmezésére. Úgy gondolom, a jövő útja az egyre nagyobb adatbázisokon alapuló, határokon átnyúló elemzéseké. Ennek alapját jelentheti a már meglévő regionális/nemzeti adatbázisok közötti adatcsere (az Applied Vegetation Science tervezett tematikus számába készülők cikkek kapcsán több ilyen projekt szervezése folyik most is), amit elősegít a vegetációs adatbázisok metaadatbázisa (<http://www.givd.info/>). Egyetértek Opponensemvel, hogy a genomikához hasonlóan, az ökológiában is a jövő fontos kutatási eszközei lesznek a nagy adatbázisok. Ezt a szükségletet hivatott európai szinten kielégíteni a LifeWatch program, amelyet az uniós nagy kutatási infrastruktúrákhoz (pl. CERN) hasonló formában finanszíroznak és működtetnek.

Opponensem kiemelten fontosnak tartja a preferenciális felvételek statisztikai elemzésének lehetőségeivel foglalkozó fejezetet. Valóban, ha a random mintavétel feltételét túl szigorúan értelmezzük, akkor az adatbázisok vizsgálati lehetőségei erősen korlátozottak lennének, és a durva skálájú ökológiai vizsgálatok esetén megfelelő számú, független ismétlés hiányában nem alkalmazhatnánk statisztikai teszteket. A megoldás a statisztikai tesztek feltételeinek merev számonkérése helyett, azok megértése, értő alkalmazása és az eredmények megfelelő interpretációja.

Az ehhez a kérdéskörhöz kapcsolódó esettanulmányban azt feltételeztük, hogy a preferenciális mintavételnél nagyobb lesz a kvadrátonkénti átlagos fajszám. Ez a hipotézis azon a közvélekedésen alapul, hogy a preferenciális mintavétel elkerüli a fajszegény foltokat. Opponensem rámutatott, hogy egy ezzel ellentétes hatás is van: az átmeneti vagy gyomos állományok fajszáma magasabb lehet,

mint a tipikusaké. A két hatás eredője az is lehet, mint azt mi is megfigyeltük, hogy az átlagos fajszám a preferenciális felvételekben kisebb.

A klasszifikáció előtt „zajsűrés” céljából alkalmazott ordináció nem a saját ötletem, bár a módszert módosítottam (az eredeti korrespondencia analízist a távolságfüggvények közötti választást lehetővé tevő főkoordináta elemzésre cseréltem) és hasznosságát elsőként vizsgáltam szimulált adatokon. A „zajsűrés” elnevezést Gauch „Noise reduction by eigenvector ordinations” című cikkéből vettem, de elfogadom Opponensem véleményét, hogy helyesebb lenne lényegkiemelésről beszélni. A metrikus sokdimenziós skálázás választása szintén ennek a cikknek az „öröksége”; mivel az eredeti cikkben is metrikus módszerekről volt szó, én is ilyenekben gondolkoztam. Érdekes lenne párhuzamos elemzéseket végezni ugyanazon az adatsoron metrikus és nem-metrikus ordinációs módszerekkel.

Válaszaim a bíráló végén feltett kérdésekre:

1. Egyes típusok felülreprezentáltsága okozhatna problémát az elemzésekben, de az adatok előzetes sztratifikációjával a „felesleges” felvételek kiszűrhetők az elemzés előtt. Erre a célra többféle módszer is ismert, én a legjobbnak Lengyel Attila nemrég publikált módszerét tartom, amely a fajösszetétel alapján végzi a felvételek kiválasztását.
2. Tradicionálisan az angolszász országokban erős hagyományai voltak a populáció-centrikus megközelítéseknek, amelyekben a társulás szintű jelenségeket is erről az alacsonyabb szerveződési szintről kiindulva próbálják megmagyarázni. Ezzel szemben a cönológiai hagyományokból kinövő kontinentális vegetációtudomány számára a közösség emergens tulajdonságokkal rendelkező, önálló szerveződési szint, amelynek működése nem érthető meg komponensek vizsgálatával. A redukcionista angolszász megközelítés számos alapvető szabály felfedezésére vezetett, amelyek közül a legfontosabbként korlátozott hasonlóság elvét emelném ki. Ezek az eredmények az európai kutatók gondolkodására is erősen hatottak. Ugyanakkor fokozatosan kiderült, hogy a sokáig lenézett és tudománytalannak tartott vegetációtípusizálás mennyire fontos egyrészt a természetvédelmi gyakorlat számára, másrészt a lokális kísérleti eredmények érvényességi körének megállapításához. Ezen a területen az angolszász világ tanulhat közép-európától! Nem véletlen, hogy Dél-Afrika nemrég elkészült és Ausztrália most készülő vegetáció-felmérését is európai kutató, Laco Mucina irányítja. Arra, hogy a vezető amerikai kutatók is felismerték a vegetáció-felmérések és az ezekhez kapcsolódó adatbázisok fontosságát, jó példa Robert Peet, az Ecological Society of America korábbi vezetője és az Ecology korábbi főszerkesztője, aki most a vegetációs adatbázisok meta-adatbázisát felügyelő bizottság egyik tagja.
3. Ha a figyelembe veendő ordinációs tengelyek számát randomizációval határoznánk meg az lassúbb, de pontosabb módszer lenne, mint a törött pálca eloszlással való összehasonlítás. A törött pálca eloszlás olyan idealizált, egyenletes eloszlás, amely random adatmátrixokban sem valósulhat meg, ha fajszámok és a fajok frekvenciái rögzítettek. Ezért azt várom, hogy a törött pálca eloszlást alkalmazva túlbecsüljük a fontos tengelyek számát, ami ronthatja az információsűrítés hatékonyságát. Az, hogy ez a túlbecslés mekkora, függ az adatsor tulajdonságaitól. A dolgozatban is szereplő sziklagyepi adatsor esetén a törött pálca eloszlás alapján a fontos tengelyek száma 9 volt. Ha a fajszámokat és a fajok frekvenciáit is megőrző keveréssel (Miklós & Podani 2004 – Ecology 85, 86-92) létrehozott random mátrixok elemzésének eredményeit tekintjük referenciának, a fontos tengelyek száma csak nyolc.

Ebben az esetben tehát a különbség nem jelentős. A hiba mértékétől függetlenül a túlbecslés kevésbé veszélyes, amit az alulbecslés, mert az utóbbinál érdemi információ veszhet el, amikor fontos tengelyeket figyelmen kívül hagyunk.

4. A klasszifikáció jóságát értékelő módszerek jelentős része azon a feltételezésen alapul, hogy a távolságmátrixban szereplő értékek tökéletesen leírják a felvételek közötti különbségeket. Azonban a megfelelő távolságképlet kiválasztása egyáltalán nem egyszerű feladat, és nem minden esetben egyértelmű, hogy a kiválasztott képlet a legjobb. Ezért elsősorban azokat a módszereket ajánlom, amelyeknél nincs szükség erre a feltevésre. Ha elegendő adatunk van, akkor mindenképpen érdemes megvizsgálni, hogy melyek azok a csoportok, amelyek több párhuzamos adatsorban is megjelennek. Ha erre nincs mód, akkor a csoportok értelmezhetőségét mérő módszereket ajánlom, ezek közül is elsősorban a karakterjajok számát, amely akkor is jó eredményt ad, ha nagyon sok a ritka faj az adatsorban.

Végezetül Opponensem munkáját még egyszer megköszönve tisztelettel kérem válaszaim elfogadását.