

BÍRÁLAT
Dr. SZABÓ SZILÁRD
TÁJMETRIAI MÓDSZEREK KRITIKAI ALKALMAZÁSA A TÁJANALÍZISBEN
című
MTA doktori disszertációjáról

1 Vélemény a dolgozat felépítéséről, stílusáról, arányairól, eredetiségéről

A szerző dolgozatában a tájmetriai metrikák kapcsolatrendszerének kritikai elemzésével kívánja vizsgálni e metrikák redundanciáját és sokféle kapcsolatrendszerét az elemzések három szintjén: foltok, az osztály és a táj szintjén. A dolgozat rendkívüli időszerűségét az a tény adja, hogy e metrikák száma az elmúlt években folyamatosan nőtt anélkül, hogy szisztematikus összehasonlításuk megtörtént volna.

Jelen bírálat írója le kívánja szögezni, hogy geológus és matematikus révén, akinek szakterülete a geostatisztika, a metrikák ökológiai, tájökológiai szerepéről, valamint megfelelőségükről a tájmetriai értelmezésekben nem kíván véleményt formálni. Ennek következtében jelen bírálat pusztán a dolgozat logikai felépítésére és az egyes összehasonlításokban alkalmazott statisztikai eljárásokra fog korlátozódni azzal a geostatisztikai szemlélettel, amely szintén a térbeli jelenségek modellezésére helyezi a hangsúlyt.

A dolgozat a bevezetőben megjelölt 11 célt 100 oldalon, 44 ábra és 28 táblázat segítségével dolgozza fel. Irodalomjegyzéke rendkívül részletes: a 263 hivatkozott tételben a szerző a tájmetriai módszerek alkalmazásának széles skáláját öleli át. (A bíráló meglepődve olvasta az ezek geológiai alkalmazási lehetőségeit taglaló hivatkozást.) A dolgozat szövegtörzsében minden irodalmi tételre történt hivatkozás. A dolgozat stílusa remek, még a kívülálló számára is olvasmányos, a fogalmazás szabatos, tömör, de mindvégig érthető. A dolgozat egyik nagy formai pozitívuma, hogy elütéseket, a képletekben index kimaradásokat nem tartalmaz.

A munka felépítése tökéletesen megfelel a tudományos munkák szerkezetének: a bevezető fejezetben világos definíciója szerepel a tájmetriának és a vizsgált metrikáknak. Az „Anyag és módszer” című fejezetben a vizsgálatokban alkalmazott eljárások áttekintése történik metrikáknaként. Az „Eredmények” fejezetben a vizsgált módszerekkel kapott eredmények bemutatása és értékelése kap helyet. Az utolsó, „Következtetések” című fejezet lényegében a kapott eredmények diszkussziója. A fejezetek terjedelmi arányai általában megfelelőek. Talán a Bevezetés szakirodalmi áttekintésének 30 oldala tűnhet kissé soknak. Ennek tömörítésével a diszkusszióra („Következtetések” fejezet) bővebb terjedelem állt volna rendelkezésre, amely a munka értékét tovább növelte volna.

A dolgozat egyik legfőbb tudományos értéke, hogy statisztikai eljárásokkal próbálja rendszerezni az idők során gyorsan és helyenként kritikátlanul elszaporodott tájmetrikákat. Kiemelkedő érdeme az objektivitás, és az a tény, hogy a gyakorlati példák mellett elvi alapon konstruált teszt területen is értékeli a metrikákat. Ez utóbbi különösen fontos a kapcsolatok objektív (mintától független) mérlegelésére. A dolgozat megközelítésének újszerűségét, a témaválasztás mellett, a vizsgálatokban használt módszerek komplexitása adja. Külön érdemes hangsúlyozni, hogy a munkában alkalmazott statisztikai eljárások végrehajtása precíz, a protokollnak megfelelő. Ahol nem vagy „nem annyira”, azok a „Tételes megjegyzésekben” képezik a kritika tárgyát.

A dolgozatban kifejtett gondolatok és feldolgozott adatok, a szerző irodalmi munkásságának hivatkozott publikációi alapján, teljesen nyilvánvalóan a szerző sajátjai.

2 Tételes észrevételek

2.1 Az elméleti foltmintázat értékelésének szubjektív jellege

Az „Anyag és módszer” fejezet bemutatja, a tájfragmentáció mérőszámainak vizsgálatához kialakított 12 változóból álló elméleti foltmintázatot. A foltmintázat értékelése a fragmentáció mérőszámainak tükrében a 4.2. fejezetben olvasható. Itt a szerző részletesen elemezte a foltmintázatok összehasonlítását a kategóriák száma és a pixelek aggregáltsága szerint. Ennek eredményét a 26. ábra foglalja össze. Az ábra rendkívül szemléletes. Sajnos értékelése pusztán empirikus úton történik, ami helyenként szubjektív lehet. Ugyanakkor a 14-16 táblázatok adatai alapján ennek egy bizonyos fokig objektívebb értékelése akár a neurális hálók, akár bármely cluster analízis vagy nem-lineáris síkra vetítés alkalmazásával megtörténhetett volna. Ennek elmaradása az olvasó számára kis hiányérzetet jelent, hiszen az objektív „elvi” mintázatokra kapott válasz-metrikák vélhetően a közölteknél több általánosítható megállapításra lennének alkalmasak.

Úgy érzem, hogy az objektív általánosítás helyett választott szubjektív értékelés ennek a szép modellnek a súlyát kissé csökkenti. Annál is inkább, ennek a fejezetnek összefoglaló értékelése a lehetőségekhez képest kissé visszafogott: „...nem lehet egyetlen metrikára alapozni a vizsgálatokban szereplő indexeket.”.

2.2 Főkomponens analízis kontra faktor analízis problémája.

Mindkettő célja a dimenziószám csökkentése. Ugyanakkor lényeges különbség a két eljárás között, hogy a főkomponens analízis során bármely adott változó varianciát teljes egészében leírjuk a főkomponensekkel, ezzel szemben a faktor analízisek során bármelyik adott változó szórásnégyzetének csak egy bizonyos részét írjuk le a faktorokkal. A faktor analízisek során bármely változó teljes varianciáját három tag összegére bontjuk: (közös variancia+egyedi variancia+hiba variancia). A „közös variancia” az adott változó varianciájának az a része, amelyet a faktorokkal leírunk. A faktor analízisek bármely implementációja során csak ezzel a közös résszel foglalkozunk. Ezt nevezzük kommunalitásnak.

A fentiekből lényegében az következik, hogy a főkomponens analízisben az input változók esetében nem engedünk meg olyan változékonyságot, amely akár a mérési bizonytalanságra, akár a változó saját bizonytalanságára utal. A faktor analízis során viszont komolyan számba vesszük a varianciának ezt a többi változóval nem magyarázható részét.

Regionalizált (térben mért) tulajdonságok esetében, és a tájmetrikák is vélhetően ide tartoznak, az ilyen bizonytalanságok figyelembe vétele erősen ajánlott. Ekkor tehát célszerű a faktor analízisek valamely implementációját választani. Az eredményként kapott faktor súly mátrix természetesen különbözni fog a főkomponens mátrixtól. Az persze nehezen képzelhető el, hogy rögzített változók esetén a bármelyik faktorban nagy faktor súllyal jelenjen meg olyan változó, amely az azonos sorszámú főkomponensben nagyon kis főkomponens súllyal van képviselve. Ugyanakkor teljes joggal várható, hogy egy adott faktorban a változók „fontossági sorrendje” más lesz, mint az azonos sorszámú főkomponensben.

A fentiek kapcsán kérdésem: a szerző miért nem vette figyelembe a mértékekhez kapcsolódó fentiekben körvonalazott bizonytalanságot?

2.3 A kumulált sajátérték százalékok információjának szükségessége

A szerző világosan megjelöli, hogy értelmezett főkomponensnek az 1 –nél nem kisebb sajátértékű mesterséges változókat tekinti. Ez teljesen rendjén való, széles körben elterjedt szabály (a scree-plottal együtt). Ugyanakkor nagyon fontos lenne látni, hogy az értelmezett főkomponensek a teljes variancia hány százalékát magyarázzák. Ez a mutató, *a kumulált sajátérték százalék, amelynek feltüntetése mind a táblázatokban, mind a szövegben következetesen hiányzik*. Ugyanakkor ez mutat rá arra, hogy a szerző által értelmezett korrelációs struktúra bizonytalansága mekkora.

2.4 Rotált és rotálatlan főkomponens mátrix.

A főkomponens mátrix oszlopainak rotációja a becslési eljárások matematikai elemzését segíti. Ennek ára, hogy az így kialakított rotált faktorok értelmezése sok alkalommal problémás lehet. A szerző által választott varimax rotáció alapvető célja, hogy minél több nullához közeli főkomponens súlyt alakítson ki. Ennek során azon változók száma, amelyekhez sok faktor szerepel nagy súllyal, kevés lesz. Azaz komoly törekvés van az „egy változó-egy rotált faktor” típusú rendszer kialakítására.

Az ugyancsak gyakran alkalmazott quartimax rotáció ezzel szemben a magyarázó faktorok számát minimalizálja. Vagyis ez a megoldás arra törekszik, hogy egy rotált faktorban minél több eredeti változót vonjon össze.

Kérdésem az, hogy a fentiek tükrében mi indokolta a varimax rotáció választását a feldolgozásban?

Tény ugyanakkor, hogy bármelyik rotációs megoldást választjuk, mindenképpen egyfajta „sugalmazott” változócsoportokat kapunk a rotált főkomponensekben. Ezzel szemben a rotálatlan megoldás azt az eredeti korrelációs struktúrát mutatja, amelynek feltárása a szerző eredeti célkitűzése volt. Az eredmények értékeléséhez és a szerző értelmezésének követéséhez hasznos kiegészítő lett volna a rotálatlan kapcsolatrendszer bemutatása.

Kérdésem, hogy mi indokolta a rotált főkomponens mátrix értelmezését, miért vetette el a szerző a (természetesen több kölcsönhatást tartalmazó) rotálatlan korrelációs struktúra értékelését?

2.5 A térbeli folytonosságú mintázatok statisztikai értékelés

A szerző a térbeli folytonosságú mintázatok értékelését egyebek között a Wilcoxon próbával és diszkriminancia analízissel végezte el. A választás tökéletes, hacsak eltekintünk a dolgozatban szereplő alcímtől: „A térbeli folytonosságú mintázatok...”. A Wilcoxon próba (mint minden „klasszikus hipotézisvizsgálat”) látókörébe ugyanis olyan minták tartoznak, amelyben az elemek között autokorrelációs kapcsolat nem lehet. Az elemeknek függetlenek kell lenniük egymástól. A földrajzi tér jelenségei ezt a feltételt általában nem elégítik ki. Pl. a zivatargócoktól távolodva a csapadék mennyisége fokozatosan csökken, a hőmérséklet fokozatosan nő, a szikes foltok nagysága és gyakorisága a szikesedés „központjától” távolodva fokozatosan csökken, vannak magas és alacsony talajszennyezettségű zónák, stb. Mindez azt jelenti, hogy ha ezek jellemzésére valamiféle mértéket használunk, akkor ennek a mértéknek mérőszámai között egy változó erősségű, de vélhetően szignifikáns autokorreláció van. Ez az a jelenség, amely a klasszikus statisztikai próbák alkalmazását ebből a körből kizárja. Az ilyen térbeli kapcsolatú folyamatok matematikai leírása a fentiek miatt az autokorrelációt, mint a térbeli folytonosság mértékét felhasználó sztochasztikus, geostatistikai elemzések középpontjában van. Lényegében ugyanez a kétség a diszkriminancia analízis elvi alkalmazhatóságával kapcsolatban is.

A fentiek kapcsán kérdésem: vajon a tájmetriai mintázatok térbeli folytonossága hogyan értendő?

2.6 A geometriai felbontás hatása a korrelációs struktúrára

A szerző 4.1. fejezetben összegzi az ezzel kapcsolatos vizsgálati eredményeit és következtetéseit. Itt a legelső bevezető gondolata a következő: „A geometriai felbontásnak kisebb hatása volt a korrelációs struktúrára, mint az várható lett volna.” Nem tudom, hogy mekkora hatás lett volna várható. Egy mérték térfogati hatásának csökkenésével ugyanis egészen biztosan nő a szórásnégyzet, aminek következtében a korreláció csökkenni fog, ha feltételezzük, hogy a jelenséget a gyenge stacionaritás alapmodelljével írjuk le, mint ahogy a szerző is tette. A korrelációs struktúra „szignifikáns” változása (a fogalmat a szerző vizsgálataiban értelmezte) csak akkor valószínű, ha a felbontás csökkenésével a vizsgált jelenség olyan léptékének mintázása történik, amelyben a térbeli eloszlást más heterogenitású folyamatok alakították ki, mint az induló durva léptékben. A fentiek miatt teljesen érthető, ha nem a legfinomabb és legdurvább felbontások korrelációs struktúrája különbözik a legjobban.

A szerző egy – bár nem bőre szabott -- alfejezetet szentel annak elemzésére, hogyan változik a korrelációs struktúra, ha a változók egy részét újakra cseréli. Ez a megfontolás a többi fejezet kifejezetten fajsúlyos megállapításaihoz képest a trivialis körbejárásának tűnhet. Valóban, hiszen egy PCA-val feltárt korrelációs struktúra minden olyan esetben változni fog, ha az input korrelációs mátrix változik. Ez pedig minden olyan esetben változni fog, ha az eredeti változók közötti feed-back rendszerekbe tartozó változók közül néhányat kicserélünk. A változás nem lesz ilyen látványos, ha nem ilyen „lényeges változók” cseréjét végezzük el.

2.7 Az eredmények extrapolálhatóságának módszertani fejlesztése

A címben szereplő extrapolálhatóság fogalom a disszertáció szöveggörnyezetében teljesen félreérthető. Az olvasó ugyanis azonnal matematikai extrapolálhatóságra gondol, és ez komoly zavart jelenthet számára, hiszen a fejezetben alkalmazott eljárások numerikus extrapolációja egyszerűen nem lehetséges. A szerző szerencsére a 4.1.5 fejezet második mondatában azonnal tisztázza, hogy itt lényegében az a kérdés, hogy a PCA-val kapott eredmények mennyire függetlenek az éppen feldogozott mintától. Ugyanakkor kár ilyen szakmailag félreérthető fogalmakkal megzavarni az egyébként tiszta gondolatmenetet.

A továbbiakban a szerző PCA értelmezett eredményeinek mintaterülettől való függetlenségét úgy igyekszik megközelíteni, hogy más, a korábbiaktól teljesen eltérő területeket vizsgál. Sajnos ez egy logikai bukfencnek tűnik. A főkomponens/faktoranalízisek értékelésének minden esetben azzal kellene zárulni, hogy a főkomponenseket/faktorokat a bennük jelentős súllyal szereplő változók által szakmailag értelmezzük. Ebből persze az is következik, hogy *a PCA mintaterülettől való függetlensége csak a főkomponensekben értelmezett jelenségek általánosíthatóságának elemzésével végezhető el és nem új mintákon történő elemzéssel.*

2.8 Keresztkorreláció fogalma

A szerző a disszertációban következetesen használja a keresztkorreláció fogalmát két változó közötti kapcsolat numerikus kifejezésére. A statisztikai gyakorlatban a keresztkorreláció fogalmát két változó idősorának összehasonlítására használják. Ezzel lényegében azt vizsgálják, hogy az idősorokban az egyik változó értékeinek valamilyen irányú elmozdulását mekkora időbeli késéssel és ekkor mekkora linearitással követi a másik változó.

A szerző ellenben a keresztkorreláció fogalmát láthatóan két, nem idősorban mért, változó közötti korreláció szinonimájaként használja. Ez komoly félreértésekhez vezet és korrigálása feltétlenül szükségesnek tűnik.

2.9 A korrelogram fogalma

Ugyancsak zavaróan hat a korrelogram fogalom használata. Félreérthetőségének alapja az, hogy a statisztikai gyakorlatban a korrelogram fogalmat az autokorrelációs együttható eltolástól függésének megjelenítésére hivatott ábra megnevezésére használják. Ilyenképpen az idősorok elemzésének egyik nagyon fontos eszköze.

Úgy tűnik, hogy a szerző a hivatkozott dolgozat (Wright, 2012) nyomán, a korrelációs mátrix szimbolikus megjelenítésére használta ezt a nevet. Ugyanakkor tény, hogy egy „klasszikus” statisztikai elemzésekből álló környezetben az ilyen fogalom használata teljesen félreérthető. Emiatt korrigálása feltétlenül szükségesnek tűnik.

2.10 A korrelációs mátrixok robusztusságának kérdése

A szerző az PCA mintázatok általánosíthatóságának elemzésével foglalkozó fejezetben a következőket állítja: „...a korrelációs mátrixok alkalmazása egy hatékony és robusztus eszköz lehet a tájanalízisben.”

Az állítás első problémája az, hogy azt sugallja: a korrelációs együtthatók (a korrelációs mátrix elemei) robusztusak. Sajnos ez így általánosságban nem igaz. Hacsak nem rangkorrelációt használunk, ezek a mértékek általában nem robusztusak, hiszen a kiütő értékekre éppúgy, mint a linearitástól való eltérésre különösen érzékenyek.

A második probléma, hogy az olvasónak az a benyomása, hogy a korrelációs mátrixok és a PCA biplot-diagramjai mintegy a főkomponensek nem-triviális értelmezhetőségét lennének hivatva ellensúlyozni. Ha valóban ez a szerző gondolatmenete, akkor a következő kérdésre feltétlenül választ kellene adnia: Miért fogadja el a korrelációs mátrixot informatívabbnak, annak tükrében, hogy a PCA ennek szerveződését mutatja be? Ha nem ez a szerző gondolatmenete, akkor viszont az idézett állítást alaposan át kellene fogalmazni.

2.11 A foltkapcsolati modellek kapcsolatainak vizsgálata cluster analízissel

A cluster analízisek az automatikus csoportosító eljárások nagyon népszerű csoportját képezik. Tény, hogy a mintatér pontjainak (Q-típusú eljárások) és a mintákon mért változók csoportjainak (R-típusú eljárások) felderítésére egyaránt alkalmasak. A vizsgálatokat a szerző a hierarchikus csoportosító eljárásokra alapozta.

Az ilyen eljárások eredménye lényegében két hasonlósági mértéken alapul. Az egyik a szorosabb értelemben vett hasonlósági mutató, amely a minták összehasonlítására alkalmas. Ez határozza meg azokat a „magokat”, amelyek köré a csoportok (clusterek) kiépülnek. A másik mutató a már definiált clusterek és minták, valamint a clusterek egymás közötti hasonlóságát definiálja. Ez dönt a minta és cluster valamint a cluster és cluster összevonásáról. Ezt szokták redukciós mértéknek nevezni. A szerző által alkalmazott megközelítésben ez utóbbi a Ward-algoritmus csoport homogenizálása volt.

Mi volt a minták közötti hasonlósági mutató? A kérdés megválaszolása azért is fontos, mert határozza meg azokat a „magokat”, amelyek mentén a dendogram felépül.

A cluster eljárásoknak nagyon sokféle implementációja van, amely egy adott mintára nézve nem feltétlenül (sőt rendszerint nem is) azonos csoportosításokat ad.

Történt e vizsgálat arra nézve, hogy más algoritmus választása mennyire változtatta volna meg a Ward-algoritmussal kapott csoportokat?

Mi volt a szakmai alapja a Ward-algoritmus választásának, és mennyire általánosítható a kapott csoportosítás?

2.12 A metrikák egyediségének vizsgálata diszkriminancia analízissel

Klasszikus értelemben a diszkriminancia analízis azt vizsgálja, hogy egy adott csoportosításban az egyes változók milyen szerepet játszanak. Ennek során az input változóknak olyan lineáris kombinációját állítjuk elő, amely a megadott csoportokat a legjobban elkülöníti.

A gondolatmenetet követve állítható, hogy a csoportosításhoz rendelhető „lényeges” változók azok, amelyek a csoportok kialakításában lényeges szerepet játszottak, a többi elhanyagolható. A szerzőnek ez az elgondolása igazolható és valódi tartalommal bír.

Ugyanakkor problémát jelenthet, hogy *ha a standardizált függvények mellet a diszkrimináló függvényekhez való változó-hozzájárulást korrelációval fejezzük ki, akkor ezek a korrelációs együtthatók a főkomponens elemzés struktúra mátrixához hasonló eredményre vezetnek. Ilyen formában ez az elemzés vélhetően nem adhat a struktúra mátrixtól szignifikánsan különböző eredményt.*

3 A TÉZISEKRŐL

A 11 tézisről, illetve az 1 tézis-csoport „a”, „b”, „c” és „d” pontjairól, az alábbi álláspontom alakult ki.

Elfogadom az 1/b, 1/c, 2., 3., 4. 5., 6., 7., 8., 9., 10. 11. számú téziseket.

Nem fogadom el tézisként az 1/a, és 1/d pontban tett állításokat. Ennek indokai a tételes észrevételekben vannak részletezve.

4 ÖSSZEGZŐ MEGÁLLAPÍTÁS

Mivel a „Tételes észrevételek” fejezetben kritizált gondolatok és feltett kérdések a disszertáció alapvető megállapításait érdemben (szignifikánsan) nem befolyásolják **kijelentem, hogy a disszertációt nyilvános védésre (tökéletesen) alkalmasnak tartom.**

Szeged, 2015. május 5.

.....

Geiger János, PhD