

Válasz Dr. Tusnádý Gábor bírálata, melyet a „Különböző taxonok pollenjeinek komplex statisztikai elemzése a meteorológiai elemekkel összefüggésben, különös tekintettel a parlagfű pollenjére” c. MTA doktori értekezésemre készített.

Igen Tisztelt Bíráló!

A bírálata a válaszat a megjelölt blokkok szerint adom meg.

Vélemény

Makra László Különböző taxonok pollenjeinek komplex statisztikai elemzése a meteorológiai elemekkel összefüggésben, különös tekintettel a parlagfű pollenjére

című MTA Doktori Értekezéséről

Az értekezés a Szeged városát veszélyeztető pollenekre vonatkozó adatsorokat dolgoz fel meteorológiai és egészségügyi adatokkal kiegészítve. Napjainkban az alkalmazott tudományokra jellemző a munkában felhasznált különböző tudományágak nagy száma. Kétféle stratégia lehetséges: nagy csoportok együttműködése avagy polihisztorok munkája. Jelen dolgozat jellegét tekintve inkább a második kategóriába sorolható, a szerző igen dicséretes módon minden érintett területen otthonosan mozog és nagy önállóságra tett szert. Végső következtetései meggyőzőek és hasznosak. (De nem idegen a szerzőtől az első stratégia sem hiszen cikkeinek többségét 3-4 társszerzővel írja.) Ezért azt javaslom, hogy az MTA fogadja el az értekezést és adja oda a szerzőnek a tudományok doktora címet.

Az alábbiakban olyan észrevételeket közlök, amelyeket kizárólag matematikai statisztikai szempontból írok és nem csökkenthetik a disszertáció értékeit.

Kiindulok a 24. ábrából (75. oldal). Ha jól értem, itt tíz év áttekintő adatsora látható valamilyen átdolgozásban, de ennek szerintem most nincs jelentősége. Apró megjegyzés, hogy ezt az ábrát jobban szeretném egész oldalon látni esetleg színesben, de nekem most a céljaimnak így is megfelel. Amit nekem ez az ábra mond az egyértelműen az, hogy egy ilyen jellegű adatsorra a klasszikus matematikai statisztika módszerei nem alkalmazhatóak. Hiszen nincs két összehasonlítható adat. Az évek mind mások, az éveken belül természetes ciklusok vannak. Valamit ilyenkor ki kell találni. Ezt természetesen nem lehet egy meteorológustól elvárni, de talán helyesebb lett volna együttműködő partnereket keresni.

Hasonló hatást mutat a 11. táblázat: a vizsgált tíz évben az összes pollenszám nagy mértékben és első ránézésre össze-vissza, rendezetlenül változott, a minimum 3857, a maximum 13 854. Itt két kérdést tartok jelentősnek, amelyekre a disszertáció nem válaszol.

A PARLAGFŰ POLLENSZÁMOK JELENTŐS NAPI ÉS ÉVI INGADOZÁSAINAK A KÉRDÉSE

1) Miért ilyen nagy az ingadozás? Laikusként azt gondolom, az általános agrár tendenciák mutatkoznak itt, de el tudom képzelni azt, hogy újabb kori politikánk összevisszasága is hatást gyakorol. A jelen támogatási rendszerben sokszor bevetik a gazdák a földet de nem várják ki a termést, hanem korábban leveszik a kultúrnövényt, ami visszahathat a gyomnövényekre.

A Kárpát-medencében a legnagyobbak a parlagfű napi, valamint évi pollenszámai. Szegeden nagyobbak az átlagok és nagyobb a pollenkoncentráció szórása is, összehasonlítva például a Kárpát-medencén kívüli két legnagyobb európai parlagfű élőhely (Lombardia, Olaszország és a Rhône-völgye, Franciaország) pollenszórásaival. Jól tükrözi ezt, hogy a napi pollenkoncentrációk gamma, vagy lognormális eloszlással közelíthetők (Matyasovszky és Makra, 2011). Ez utóbbi élőhelyeken kevesebb az ugar és a betakarítás után hamar megtörténik a tarlóhántás, amelyek kisebb pollenszámokkal járnak. A professzor úr által mondottakkal összhangban, az évi összes parlagfű pollenszám nagy ingadozásának okai lehetnek az agrárpolitika változó preferenciái. Ezek egyik következménye lehet a fiatal ugaron hagyott területek (mint korábbi fontos élőhely) eltűnése, ami a pollenszámok csökkenésével jár. Ennek okai közé sorolhatók többek között a legeltetés, erdősítés, beépítés, s a rendszeres kaszálás. Ugyanakkor, ha nagyobb az ugar, valamint az ember által elhagyott területek kiterjedése nő a tájban, akkor a parlagfű tovább terjedhet, különösen a homoktalajokon. További befolyásoló tényező az évről évre változó időjárás. Ezzel összefüggésben – különösen meleg és száraz tavaszokat követően – a termés korábbi betakarítása miatt a tarlóhántásig hosszabb idő telik el. Ez a parlagfű nagyobb elterjedéséhez és a pollenszámok növekedéséhez vezethet. A parlagfű jól alkalmazkodott a meleg és száraz klímához. Ugyanakkor extrém meleg időjárási feltételek (amelyek alacsony csapadékmennyiséggel járnak) korlátozhatják a parlagfűpollen koncentrációt. Ugyanis ekkor a növény a túlélésre törekszik (arra, hogy megőrizze az életfolyamataihoz szükséges vizet a szervezetében), semmint a terjeszkedésre. Másrészt, ha a nyári hónapok csapadékösszege magas, az megnöveli a biomasszáját, ami nagyobb pollenszóráshoz vezet.

Hivatkozás

Matyasovszky, I., Makra, L., 2011. Autoregressive modelling of daily ragweed pollen concentrations for Szeged in Hungary. *Theoretical and Applied Climatology*, 104(1-2), 277-283.

A NAPI POLLENKONCENTRÁCIÓK MEGMARADÁSI HAJLAMA

2) Hogyan lehet ezt az igen jelentős hatást a statisztikai feldolgozásban kompenzálni? A 10. fejezet rövidtávú (1-3 napos) előrejelzéssel foglalkozik. Szerintem ezek hatékonyságát elsősorban ez a globális ingadozás magyarázza. Bő termés esetén is szűk termés esetén is a szomszédos napok egyformának mutatkoznak.

Igen, az előző napi pollenkoncentráció mindig fontos magyarázó változó volt. És valóban, van egy megmaradási hajlam a napi pollenkoncentrációkban, azaz a mai pollenkoncentráció nem tér el jelentősen a tegnaptól, azonban ezt befolyásolja az időjárás módosító hatása. Ugyanakkor a „Bő termés esetén is, szűk termés esetén is a szomszédos napok egyformának mutatkoznak” némiképp leegyszerűsítő magyarázat, mert a napi koncentrációk ingadozása igen jelentős lehet, amit megpróbálhatunk a meteorológiai elemek ingadozásával kapcsolatba hozni (Matyasovszky et al., 2012).

Hivatkozás

Matyasovszky, I., Makra, L., Csépe, Z., 2012. Associations between weather conditions and ragweed pollen variations in Szeged, Hungary. *Arhiv Za Higijenu Rada I Toksikologiju*, 63(3), 311-320.

A VÁLTOZÓ POLLENKONCENTRÁCIÓJÚ ÉVEK ÉS A SZIGNIFIKÁNS ELTÉRÉS A MAKRA PRÓBA ALKALMAZÁSAKOR

Most visszatérek a 11. ábrára (36. oldal). Szemmel láthatóan ez az egyik fő eredménye a dolgozatnak. Az első észrevételem az, hogy egy ideje már töprengem azon, mi is lehetett itt a kérdés, amire nagyon bonyolult eljárással a szerző választ keresett, de nem tudom kitalálni. Az eredményt értem: augusztus 18 és szeptember 13 között a legmagasabb a parlagfű pollenje

Szegeden. Csak hát ezt nem hiszem el, mert mint láttuk, minden év más, és nem tudom, mit jelent az, hogy szignifikánsan eltérő.

Igen, minden év más, ez a meteorológiában is mindig így van. Éppen ezért kell több évet (a lehető legtöbbet nézni), hogy ne egyetlen évre jellemző eredményt kapjunk, hanem a teljes adathalmazra (ebben az esetben 22 év) lehessen következtetést levonni. Egyébként, ha az éghajlatváltozástól (földhasználat változástól, stb.) eltekintünk, az egyes évek véletlenszerűen és nem szisztematikusan különböznek. A véletlenszerű változás viszont benne van a Makra-próba próbastatisztikájának nevezőjében (itt van a becsült szórás).

A MAKRA PRÓBA HATÉKONYSÁGÁNAK A KÉRDÉSE

Így jutunk vissza a bajok első gyökeréhez, a Makra-próbához (15. oldal). A kérdés az, hogy egy homogén adatsor valamely része eltér-e az egésztől. Ha különválasztjuk a részt az egésztől, kapjuk annak másik részét. Ezek összehasonlítására vonatkozik a kétmintás u- és t-próba, amelyben a két rész átlagának az eltérését vesszük. Legyen ez az átlag mondjuk az egyikben 100, a másikban 200, és legyenek a méretek egyformák, akkor a különbség 100 az átlagokra és 50 a nagy átlagnak az első átlagtól vett eltérésére. Tehát hatékonyságot veszünk azzal, hogy összevonjuk a két csoportot. Ha a szórásokat is figyelembe vesszük, egzaktul kimutatható az eltérés.

Igen, ez így van!

Azonban a pollenkoncentrációk évi menete annyira erős, hogy szignifikánsan eltérő alperiódusok ezzel a kevésbé hatékony próbával is bizonyosan kimutathatók. Ahogy ez ki is jött. Korábban matematikusokkal is konzultáltam, s több cikkemben alkalmaztam a módszert és a most a professzor úr által említett problémát soha nem vetették fel.

JAVASLAT A POLLENADATOK FELDOLGOZÁSÁRA

Én azzal kezdtem volna az adatok feldolgozását, hogy megkeresem a tíz év legfontosabb globális mutatóit a megvizsgált fajokra: évi összes pollenszám, annak koncentrátsága. Minden egyes fajra a napi vagy óránkénti pollenszám egy folytonos és pozitív függvény közelítései, amit az integráljával leosztva sűrűségfüggvényt kapunk. Így a valószínűségi változókra kidolgozott statisztikák (átlag, szórás, momentumok) alkalmazhatóak. Érdekes lehet a fajok közti kapcsolat, ha valamelyik szárazságtűrő másként viselkedik mint amelyik a sok csapadékot kedveli. Meg kell választani a megfelelő időszakaszt, ez talán a hét lehet, és elképzelhető, hogy a Rasch modell leírja a keletkezett tábla legfontosabb hatásait. Ez hasonló a szórásanalízisre, de pozitív mennyiségek tárgyalására alkalmas.

Igen, ha minden egyes évre elvégezzük a professzor úr által említett eljárást, az egyes évek napi pollenkoncentrációinak évi menete egy univerzális skálára kerül, és jobban összehasonlíthatóvá válik. Így valóban jobb képet kaphatunk a pollen csúcs koncentrátságáról. Ezt azonban nem terveztük vizsgálni a dolgozatban.

A különböző taxonok időjárásfüggésének a Rasch modell segítségével történő vizsgálata egy érdekes megközelítési lehetőség. Az általunk feldolgozott irodalomban egyáltalán nem talákoztunk vele, éppen ezért a további munkáinkban szívesen alkalmazzuk majd a modellt.

AZ OBJEKTÍV ÉS SZUBJEKTÍV CSOPORTOSÍTÁSSAL KAPCSOLATOS MEGJEGYZÉSEK

Túl a belső analízisen keresni kell külső globális mennyiségeket: csapadék, szél, napsütés. A mechanikus tábla módszer helyett sokkal jobb az idősoros módszer, amikor az időbeli

változásokat a múlt és a külső változók hatásaként keressük. Ez továbbfejleszthető előrejelzésre és kontrollra: annak meghatározására, hogyan lehet egyrészt a lakosságot riasztani másrészt mi a legalkalmasabb védekezés. A vizsgált növényeknek jól meghatározott szezonális ciklusa van amit nem lehet figyelmen kívül hagyni. A sztochasztikus kapcsolatok keresésének két élesen eltérő módja van, az egyik az eddig vázolt időben folytonos többváltozós analízis, a másik csoportok képzése, vagy más néven keverékek felbontása. A szerző ez utóbbit használja, amit elvileg nem lehet kifogásolni, nagyon nehéz eldönteni, mikor melyik módszer hatékonyabb. A tapasztalatom szerint jobb mind a kettőt kipróbálni, ha valami alapvető szakmai érv nem segít a döntésben. Kétféle csoportosítást használ a szerző, az egyik a széljárások adatait csoportosítja először (objektív csoportok), aztán a csoportok szerint megnézi a pollenekre gyakorolt hatást. A másik egy hazai irodalmi csoportosítás (szubjektív csoportok). Meg kellene vizsgálni ennek a kettőnek a viszonyát: minden egyes napnak van a kétfajta csoportosítás szerinti beosztása. Ezekből kell keresztábrát képezni. Az első kérdés az, hogy van-e a két csoportosítás között kapcsolat. Laikusként azt várom, hogy szoros kapcsolat legyen. Miután ez kiderült, elvileg meg lehet kérdezni, melyik csoportosítás hat jobban a pollenekre avagy a kettőt egyszerre használni. A két csoportosítás hatékonyságát összehasonlítja a szerző, de mechanikusan a szignifikáns esetek számát használva. Tekintve, hogy az objektív csoportosításban 8, a szubjektívben 13 csoport van, ez így nem helyes. A szerző szemmel láthatóan nem ismeri Gulyás Ottó munkásságát, aki szintén kidolgozott időjárási csoportokat (Ambrózy Pál, Barholy Judit és Gulyás Ottó: Évszakos bontású makroszinoptikus típusok kialakítása clusteranalízissel az atlanti-európai térségre, Országos Meteorológiai Szolgálat, 1983). Dévényi Dezsővel közösen írt könyvet Matematikai statisztikai módszerek a meteorológiában címmel.

Ha nagyobb a csoportok száma, akkor nagyobb az esély, hogy több párt talál a próba szignifikánsan különbözőnek. Ha a szignifikáns különbségeket elosztom az összes lehetséges párok számával (%), akkor ezek már összehasonlíthatók, mint ahogy a dolgozatban is történt (5. táblázat, 52. oldal; 7. táblázat, 60. oldal; 8-9. táblázat, 62. oldal).

Gulyás Ottó a munkatársaival évszakos bontású makroszinoptikus típusokat állított elő az atlanti-európai térségre, s készítettek egy katalógust, amely tartalmazza azt, hogy a vizsgált időszak napjai melyik típusba tartoznak az 1980-as évek elejéig (Ambrózy et al., 1983). Gulyás Ottóék évszakos típusokat állítottak elő, azaz minden évszakban eltérőek a típusrendszerek. Így nehéz összehasonlítani az ő típusaikat a dolgozatban előállított objektív típusokkal, melyeket az 1997-2001 közötti ötéves időszak napi meteorológiai adatai alapján készítettünk, és amik a pollinációs időszakra vonatkoznak.

Hivatkozás

Ambrózy, P, Barholy, J, Gulyás, O, 1983. Évszakos bontású makroszinoptikus típusok kialakítása clusteranalízissel az atlanti-európai térségre. Meteorológiai tanulmányok, No. 39. Országos Meteorológiai Szolgálat

A HELYI, A KÖZEPES TÁVOLSÁGÚ ÉS A NAGYTÁVOLSÁGÚ POLLENTANSZPORT KÉRDÉSE

Sokat foglalkozik a szerző azzal a kérdéssel, milyen arányban vesznek részt Szeged levegőjének szennyezésében a közeli (helyi), közepesen távoli (Kárpát-medencei) és egészen távoli (európai) források. Ez valóban fontos kérdés, hiszen nagy mértékben hatása lehet a megelőzésre. A szerző nem mond számokat, talán 80, 15 és 5 százalék lehet a véleménye. Ez valóban fontos kérdés, de szerintem a rendelkezésre álló adatok alapján erre nem lehet válaszolni. Egy lehetséges módszer a pollenek mikro-analízise, keverékfelbontása lehetne, feltehetőleg van eltérés a különböző források között.

A nagy távolságú pollen transzportnak a lokális pollenszórás is magába foglaló közepes távolságú pollentranszporttól való szétválasztására egy heurisztikus közelítést adtunk. Az eljárás alkalmazásával számszerű adatokat nem tudtunk előállítani, csupán a két komponens relatív súlyát adtuk meg a mért pollenkoncentrációban. Ugyanakkor a kapott súlyok csak a napi pollenkoncentrációknak a vizsgált magyarázó változók által megmagyarázott variancia-hányadára érvényesek, és semmit nem tudunk mondani arra a variancia-hányadra, amely a vizsgált magyarázó változókkal nem magyarázható. Az a következtetésünk, hogy Szeged esetében a közepes távolságú transzportnak nagyobb a súlya a napi pollenkoncentrációk alakításában mind a nem-csapadékos, mind pedig a csapadékos napokon, összhangban van Juhász és Juhász (1997) megállapításaival, mivel a napi csúcs pollenszámok 8,00-16,00 óra között lépnek föl (Juhász és Juhász, 1997).

Hivatkozás

Juhász, I. E., és Juhász M., 1997. Az aeroallergén növények pollenszórásának napi ritmusa a Dél-Alföld levegőjében. In: Környezeti ártalmak és a légzőrendszer. Vol. 7, (szerk: Szabó, T. és Miriszlai, E., pp. 28-34. Hévíz, Budapest

A CSAPADÉKOS ÉS NEM-CSAPADÉKOS NAPOK, SZÉTVÁLASZTÁSÁNAK A KÉRDÉSE

Külön vizsgálja a szerző a csapadékos napokat. Első hallásra engem ez kicsit zavar, mert merevnek érzem a kettévágást: laikusként egy folytonos skálát jobbnak tartok (napi összes csapadék). Itt jutunk el a következő módszertani fejezethez, a regresszió analízisre. Ebben hiányolom a legjobban az idősoros szemléletet. Tapasztalatom szerint a lépésenkénti regresszió eredményei nehezen értékelhetőek ki, ezért csak akkor alkalmazom, ha nincs más.

A vizsgált napok durván 2/3-ában a napi csapadékösszeg 0. Emiatt vizsgáltuk külön a csapadékos és a nem-csapadékos napokat. A csapadékos napokon bevettük a magyarázó változók közé a napi csapadékösszeget is, azonban ez gyakorlatilag semmi javulást nem hozott a napi pollenkoncentráció becslésében, így végül eltekintettünk tőle.

AZ AMBROSIA POLLEN IGEN MAGAS KONCENTRÁLTSÁGÁNAK A KÉRDÉSE AZ EGYIK OBJEKTÍV CLUSTERBEN

Az objektív és szubjektív csoportosításhoz a szerző 12 meteorológiai adatot használt. A légúti megbetegedések többváltozós analízisében viszont csak ötöt (hőmérséklet, szélesebbesség, relatív nedvesség, globálsugárzás és légnyomás). Apró megjegyzésem az, hogy a 12 között négy hőmérséklet van, nem derül ki, az öt között melyiket használja, feltehetőleg az átlagot. De ami sokkal lényegesebb, az az, hogy a 14a táblázat (79. oldal) szerint a 3. csoport kiugróan magas napi Ambrosia koncentrátságot mutat. Ilyet nem látok sem a 4. táblázatban (50. oldal) sem a 6. táblázatban (59. oldal). Mi lehet ennek az oka? Elképzelhető, hogy az, hogy most eleve külön lett véve az Ambrosia pollenszezonja és a pollenmentes időszak? Ha igen, ez egy alapvető módszertani hibára hívja fel a figyelmet: amint láttuk, az Ambrosia szezonja enyhén ingadozik, ezt akarhatjuk a meteorológiai adatokból kiolvasni, de határozottan intervallum jellegű időszak. Ilyesmit csoportosítással megjósolni nem helyes. A pollenszennyezettség, meteorológiai adatok és a különböző korú betegek szezonális változásai elsősorban saját belső törvényszerűségeiket követik. Ha ezt nem vesszük figyelembe, akkor nem csoda, ha a regressziós együtthatók együtthatójának még az előjele is nehezen értelmezhető.

Az objektív és a szubjektív csoportosítás esetében egyaránt (4. és 6. táblázat), ugyanazt a 12 meteorológiai elemet használtuk. Míg az objektív esetben e 12 paraméter leginkább hasonló együttes értékeit tartalmazó napok kerülnek egy csoportba, a szubjektív esetben már előre adottak a vizsgált időszak napjainak a csoportjai, melyek az Észak-atlanti –

európai térség 0 órakor mért tengerszinti légnyomási mezőinek a hasonlósága alapján lettek szubjektíven osztályozva. Az így kapott szubjektív osztályokat a hozzájuk tartozó napok 12 meteorológiai elemének átlagaival jellemezhetjük. Míg az objektív és a szubjektív osztályozásnál a meteorológiai elemek egy széles körét vettük figyelembe (4. és 6. táblázat), a 14a. táblázat esetében, amikor a légúti megbetegedések kapcsolatát vizsgáltuk a meteorológiai paraméterekkel, valamint a biológiai és kémiai légszennyezőkkel, csupán az öt legfontosabb meteorológiai elemmel számoltunk. A hőmérsékletnek és a többi meteorológiai elemnek is a napi átlaga szerepel a táblázatban. A 14a. táblázatban az *Ambrosia* pollenszezonjának a napjaira határoztuk meg a vizsgált meteorológiai elemek leginkább hasonló napi elemegyütteseinek csoportjait, majd a kapott csoportokhoz tartozó napok adatai alapján az egyes csoportok biológiai és kémiai légszennyezőinek átlagértékeit. A 14a. táblázatban a 3. csoportba azok a napok tartoznak, amelyeknek kiemelkedően magas a parlagfűpollen koncentrációja. Ugyanakkor a vizsgált időszakok eltérőek. A 4. és 6. táblázatban 24 taxont osztályoztunk az 1997-2001 közötti ötéves periódusban e taxonok február 1 – október 31 közötti egyesített pollenszezonjára, amely a legkorábbi és a legkésőbbi pollenszórású taxon pollinációs időszakát is magába foglalja. A 14a. táblázatban viszont az 1999-2007 közötti 9 éves időszakot vizsgáltuk az *Ambrosia* július 15 – október 15 közötti pollenszezonjára.

Mint mondtam, ellenvetéseim nem csorbíthatják az értekezés érdemeit. Többnyire csak lehetséges ötleteket vetettem fel, amelyek hasznossága csak akkor derülne ki, ha valaki követné azokat. Az a tapasztalatom, hogy egy komplex adatmező mondanivalóját kiolvasni nagyon sok ötlet kipróbálásával lehet csupán. Előre nem látható módon néha az egyik sikeres, a másik nem. Ez a szellem nem idegen a szerzőtől, érezhetően maga is egyes próbálkozásaival jobban meg van elégedve mint másokkal. Remélem még sokáig fog hasonló kérdéseket vizsgálni. Javaslom a nyilvános vita kitűzését és a tudományok doktora fokozat megadását.

Budapest, 2012. november 30.

Tusnády Gábor

Szeged, 2013. május 8.

Megköszönöm a Tisztelt Bíráló munkáját – észrevételeit, megjegyzéseit.

Tisztelettel

Makra László