

dc\_536\_12

**MTA DOKTORI ÉRTEKEZÉS**

**TÉZISEK**

**MATYASOVSZKY ISTVÁN**

**NÉHÁNY STATISZTIKUS MÓDSZER AZ ELMÉLETI ÉS  
ALKALMAZOTT KLIMATOLÓGIAI VIZSGÁLATOKBAN**

**BUDAPEST, 2013. JANUÁR**

## 1. Bevezetés

Az éghajlat alapvetően statisztikus természetű, ezért vizsgálata a valószínűség-számítás és matematikai statisztika eszközeit igényli. Ezekkel a kérdésekkel foglalkozik a statisztikus klimatológia. Az értekezésben modern matematikai statisztikai eszközöket mutatunk be és használunk fel elméleti és alkalmazott klimatológiai vizsgálatokban. Célunk néhány széles körben hasznosítható, ám kevésbé elterjedt módszer megismertetése, majd egy-egy alkalmazásának tárgyalása. Az önmagukban is értékes eredményekkel egyben érzékeltetjük a bennük rejlő további gyakorlati lehetőségeket.

Részben az általános megismerés, részben a jövő éghajlatának becslési lehetőségeinek értékelése szempontjából fontos a már lezajlott, illetve zajló éghajlatváltozás detektálása és becslése. Az éghajlatváltozást legegyszerűbben a várható érték időbeli változásával szokás leírni, vagyis a trendelemzéssel. Ezzel foglalkozik az értekezés első része.

Az elméleti, de talán még inkább az alkalmazott klimatológiai problémák során fontos igény a különböző változók közötti statisztikai kapcsolat feltárása. Ilyen módszerek tárgyalása és gyakorlati alkalmazása történik az újabb részben.

Az egyik legfontosabb statisztikus klimatológiai eszköz a spektrálanalízis. Az igen szerteágazó problémakör néhány speciális vetületét vizsgáljuk az értekezés harmadik részében.

Mivel az éghajlat nagyszámú nemlineáris kölcsönhatás eredményeként jön létre, célszerű az éghajlati idősorokat a jól ismert lineáris idősor modellek helyett nemlineáris idősor modellek segítségével elemezni. Egyebek mellett ilyen lehetőséget tárgyal a negyedik rész.

## 2. Adatok, számítógépes programok

A felhasznált adatok nagy része internetről letölthető; a forrásokat az értekezésben közöljük. Az adatok egy másik, kisebb részét (az értekezésben közöljük) a Szegedi Tudományegyetem Éghajlati és Tájföldrajzi Tanszéke bocsátotta rendelkezésünkre.

Elemzéseink mindegyike saját fejlesztésű számítógépes programmal készült, az Origin 6 grafikus programon kívül egyéb szoftvert nem vettünk igénybe.

## 3. Eredmények

1.

Felhasználtunk (Cline, et al., 1995; Wishart, 2009), illetve részben kidolgoztunk (Matyasovszky, 2011; Matyasovszky and Ljungqvist, 2012) egy nemparaméteres regressziós technikán (Matyasovszky, 1992; 1998) alapuló módszert, mely alkalmas a hirtelen éghajlatváltozások (abrupt climate changes) detektálására amennyiben a rendelkezésre álló adatsor elemei kellő sűrűségben állnak rendelkezésre az adatsorban mutatkozó trend simaságához képest. A jelentős módszertani újítást az Északi Hemiszféra évi középhőmérsékleti anomáliáinak az 1850-2009 évek időszakára, az Északi Hemiszféra rekonstruált évi hőmérsékleti sorára az i.sz. 200-1995 évekre, továbbá az elmúlt 11700 évre vonatkozó oxigén izotóp adatokra, pontosabban az NGRIP (North Greenland Ice Core Project) jégfurat  $O^{18}/O^{16}$  izotóparányával kapcsolatos Holocén  $\delta^{18}O$  adatokra alkalmaztuk.

Az első esetben az 1901, 1914, 1942, 1963 és 1975 éveket detektáltuk hirtelen változási időpontoknak. Az átfogó tendencia természetesen a melegedés, ám három hirtelen

hűlési időpont (1901, 1942 és 1963) is található. Korábban nem talákoztunk olyan tanulmánnyal, mely a hirtelen változások ilyen finom szerkezetét tárta volna fel.

A hosszabb időskálát vizsgálva objektív definíciót nyújtottunk a Középkori Meleg Periódus (Medieval Warm Period: MWP) és a Kis Jégkorszak (Little Ice Age: LIA) beazonosítására. Ezek a kifejezések széles körben elterjedtek, jóllehet nincs is egyértelműen elfogadott definíciójuk (Bradley et al., 2003). A hirtelen változások alapján az MWP a 795-1120 időszakra tehető, ami némiképp korábbi és hosszabb, mint Lamb (1977) időszaka. A LIA két fázisban jelentkezik: egy erősebb és hosszabb periódus 1387-1656 között, és egy gyengébb és rövidebb időszak 1749-1883 között. A LIA ezért (kerekítve) az 1390-1880 évekre tehető, ami hosszabb, mint Grove (1988) időszaka. A legnyilvánvalóbb változás azonban a 19. század végén (1883) hirtelen meginduló intenzív melegedés.

A Holocén időszak legjellemzőbb éghajlati epizódja a Holocén Éghajlati Optimum (Holocene Climate Optimum: HCO). Mivel ez a meleg periódus fokozatos lehüléssel ért véget, behatárolása korábban eléggé bizonytalan volt. A detektált hirtelen változások alapján (kerekítve) 9900-3300 évvel ezelőttre datálható, ami jóval hosszabb, mint a mások által definiált időszakok.

2.

A hirtelen éghajlatváltozásokat a még hosszabb, 100000 éves időskálán is vizsgáltuk (Matyasovszky, 2011). Az adatsor talán legjellemzőbb vonása a Dansgaard-Oeschger (DO) események jelentkezése. Egy DO esemény néhány évszázados időszakra terjed ki egy gyors felmelegedéssel és egy lényegesen lassúbb lehüléssel. Mivel az NGRIP jégfurat  $O^{18}/O^{16}$  izotóparányával kapcsolatos 122950 év hosszúságú  $\delta^{18}O$  adatok 50 éves átlagokként állnak rendelkezésre, ezért az előző pontban említett eljárás ezúttal nem volt alkalmazható.

A DO események karakterisztikus ideje és az adatok időbeli sűrűségének viszonyából fakadóan ezek az események most az adatsor varianciájának ingadozásaként jelennek meg. A variancia ilyen viselkedését egy AR-GJR-ARCH modell (Engle, 1982; Glosten et al., 1993) további finomításával írtuk le. Ennek segítségével a DO események korábbi döntően szubjektív beazonosítása helyett egy döntően objektív eljárással detektáltuk a DO eseményeket. Az általánosan elfogadott 25 (Dansgaard et al., 1993) DO eseményen túl 3 további esetet találtunk.

Megjegyezzük, hogy ARCH (autoregressive conditional heteroscedastic) modell korábbi meteorológiai alkalmazására nem találtunk példát.

3.

Az éghajlati idősorok spektrálanalízisének irodalma hihetetlenül gazdag. A folytonos spektrumot jellemző spektrális sűrűségfüggvény (a kontinuum-számoságú frekvenciák fontosságát jellemző függvény) becslése rendszerint a periodogram simításán (nemparaméteres regresszióján) alapszik. Az éghajlati adatsorok azonban általában nem mentesek a diszkrét periódusoktól (gondoljunk például az évi menetre), sőt a spektrálanalízis egyik fő feladata éppen a diszkrét periódusok detektálása. E frekvenciáknál a periodogram viselkedése alapvetően eltér a folytonos spektrumra jellemző periodogram elemek viselkedésétől. A cél tehát olyan, ún. robusztus eljárást értelmezni a spektrális sűrűségfüggvény becslésére, amely gyakorlatilag nem vesz tudomást az ilyen kiugró értékekről. A gondolat ugyan nem előzmény nélküli a meteorológiai irodalomban, ám Mann and Lee (1996) ún. medián szűrője pusztán csak segédeszköz egy elsőrendű autoregresszív (AR(1)) modellel nyert spektrális sűrűség pontosításához.

Egy, a meteorológiai irodalomban korábban nem ismert robusztus becslés (Janas and von Sachs, 1995) finomítását alkalmaztuk (Matyasovszky, 2010a) például a NAO index

(Ponta Delgada és Stykkisholmur/Reykjavik havi tengerszinti átlagos légnyomáskülönbsége) havi adatsorának 1865-2002 időszakára. Az eredmények alapján elmondhatjuk, hogy a módszer kiállja az összehasonlítást, például a rendkívül hatékonynak tartott wavelet alapú spektrálanalízissel (Nicolay et al., 2008) is.

4.

A spektrálanalízis során széles körben követett gyakorlat, hogy AR(1) modellt illesztenek az adatsorhoz, és ha a periodogram adott frekvenciánál vagy a frekvenciák egy tartományánál meghalad egy küszöböt, akkor a spektrum ebben a pontban vagy tartományon különbözik az AR(1) spektrumtól. A küszöb nyilvánvalóan függ az AR(1) spektrumtól és a választott szignifikancia-szinttől. Az eredmények interpretációjakor természetesen azok a frekvenciák a fontosak, amelyeknél a modelltől való különbözőség megjelenik.

Az éghajlati adatsorok általában ekvidisztánsan elhelyezkedő időpontokban állnak rendelkezésre. Olykor azonban ez nem teljesül, amire jó példát szolgáltatnak a paleoklíma adatok. Nem ekvidisztáns időpontok esetében mind az AR(1) modell illesztése, mind a periodogram definíciója módosításra szorul. Az első feladattal Mudelsee (2002) foglalkozott az autoregresszív együttható legkisebb négyzetes (ordinary least squares: OLS) becslésével. Kimutattuk azonban, és példákkal is alátámasztottuk, hogy ez a számos paleoklimatológiai vizsgálatban alkalmazott eljárás igen erősen túlbecsli a spektrális sűrűséget az alacsony frekvenciáknál. A probléma kiküszöbölésére az általunk javasolt súlyozott legkisebb négyzetes (weighted least squares: WLS) becslés alkalmas (Matyasovszky, 2013a). Az OLS módszeren alapuló, de az egyes frekvenciákat egyedileg kezelő Lomb-Scargle (L-S) periodogram (Lomb, 1976; Scargle, 1982) helyett az összes frekvenciát együttesen kezelő és ezért teljes legkisebb négyzetes (total least squares: TLS) becslésnek nevezett eljárást vezettünk be a periodogram előállítására (Matyasovszky, 2013a). A GISP2 Oxigén izotóp

adatoknak a 15000 - 60000 évvel ezelőtti időszakára vonatkozó WLS-AR(1) modellünk és TLS periodogramunk Schulz and Mudelsee (2002) tanulmányának több furcsaságát tárja fel és magyarázza meg. Az ő OLS-AR(1) modelljükön és az L-S periodogramon alapuló, számos paleoklimatológiai vizsgálatban alkalmazott ún. REDFIT eljárásuk ugyanis komoly pontatlanságot rejt magában. Hasonló következtetésre jutottunk Vostok deuterium tartalom (az elmúlt 422766 évben) adatsorának elemzésekor is (Matyasovszky, 2010b). Csak egyetlen példaként említjük, hogy egy, a REDFIT szerint jelentéktelen 400000 éves ciklus módszerünk szerint statisztikailag szignifikánsnak mutatkozott. Ez igen érdekes tanulság, mert az excentricitás időbeli változásai alapján a paleoklíma adatokban logikusan várható körülbelül 400000 éves periódus eddig csak igen kevés adatsorban volt kimutatható.

5.

Az éghajlati adatsorokat jelentős részben vörös zaj jellemzi, vagyis az egyre kisebb frekvenciák (egyre nagyobb periódusidők) egyre fontosabb szerepet játszanak az adatsor kialakításában. Másképp szólva, a spektrális sűrűség a zérus felé monoton növekvő. A vörös zaj spektrumot az adatsorhoz illesztett AR(1) modell spektrális sűrűségével szokás közelíteni, ami a legegyszerűbb, de egyáltalán nem biztos, hogy kielégítő eljárás. Nyilvánvaló ugyanis, hogy a zérus felé monoton növekvő nem negatív függvények osztálya jóval tágabb, mint az AR(1) spektrális sűrűség osztálya.

A vörös zaj becslésére ezért a meteorológiai irodalomban még nem alkalmazott ún. izoton regresszió (IR) módszerén alapuló eljárást javasoltuk (Matyasovszky, 2013b). Ekkor olyan görbét keresünk, mely a periodogramot négyzetes hibaösszegben optimálisan közelíti a vörös zajjal kapcsolatos monotonitási feltétel teljesülése mellett (Zhao and Woodroffe, 2012). Pontosabban, a 3. pontban már említett okból kifolyólag a feladat robusztus változatát tekintettük (Álvarez and Yohai, 2011). Az Északi Hemiszféra 200-1995 évek közötti

hőmérsékleti sorára alkalmazva az eljárást az AR(1) és az IR spektrális sűrűségek között óriási különbség mutatkozott az alacsony frekvenciáknál. Jól példázza ez azt, hogy a vörös zaj spektrumnak AR(1) modellel történő közelítése esetenként rendkívül megbízhatatlan lehet. Ez természetesen hatással van a tényleges spektrumnak a vörös zajtól való eltérésének detektálására is (Matyasovszky, 2013b).

6.

NGRIP  $O^{18}/O^{16}$  izotóparányával kapcsolatos  $\delta^{18}O$  és Vostok deutérium tartalmának 122950 év hosszúságú adatsorának együttesét vizsgáltuk egy vektorértékű nemlineáris idősor modellel (Matyasovszky, 2010c). Ez a vektor küszöbmodell vagy VTAR (vector threshold autoregressive: VTAR) modell több rezsimből álló vektor AR (VAR) modell, ahol az adott időponthoz tartozó aktuális rezsimet egy ún. küszöbváltozó múltbeli értéke határozza meg. A konkrét esetben a küszöbváltozó NGRIP vagy Vostok egyaránt lehet. Az elemzés célja az volt, hogy kiderítsük, hogy az Arktisz és Antarktisz éghajlatingadozásiban megfigyelhető fázisreláció kialakításában elsődlegesen melyik tag tölti be a vezető szerepet. A két félteke klímaingadozásainak fázisrelációját (időbeli eltolódását) magyarázó elmélet az ún. bipoláris mérleg, melyben komoly helyet foglalnak el a globális óceáni cirkuláció változásai (Severinghaus, 2009). Az nem teljesen tisztázott, hogy a mechanizmust az Északi vagy a Déli Félteke irányítja (Seidov et al., 2001), de Steig and Alley (2002) szerint ez pusztán statisztikai elemzéssel nem is deríthető ki. Ez utóbbi megállapítás igaz lehet az adatsorok hagyományos lineáris elemzésére, ám a nemlineáris modellezéssel választ adtunk a kérdésre.

Mivel a kétváltozós VTAR modell illesztése során (Tsay, 1998) kiderült, hogy a küszöbváltozó az Arktiszt reprezentáló NGRIP, ezért NGRIP a két adatsor együttes viselkedésében alapvető szerepet játszik. Annak kiderítésére, hogy NGRIP dominanciája egyértelmű-e meg kellett vizsgálni, hogy Vostok pusztán csak reagál-e NGRIP-re, vagy vissza



is hat rá, azaz inkább kölcsönhatásban állnak. Ezt a Granger-féle okozatiság (Granger, 1969) elvének alkalmazásával derítettük fel. E szerint az Arktisz-Antarktisz éghajlat-ingadozások statisztikailag kölcsönhatásban állnak, ám ebben a kölcsönhatásban inkább az Arktisznak van elsődleges szerepe. Ez abból látható, hogy a küszöbváltozó az NGRIP, illetve, hogy a VTAR modellhez tartozó zaj varianciája a VAR modellhez képest nagyobb mértékben csökken Vostok, mint NGRIP esetén. Vagyis a nemlineáris modellben (szemben a lineáris VAR modellel) NGRIP nagyobb információt nyújt Vostokra nézve, mint megfordítva.

Egy stacionárius AR folyamat egyetlen fixponttal rendelkezik. Ez úgy értendő, hogy a folyamat  $l$ -lépéses előrejelzése a rendelkezésre álló idősor konkrét értékeitől függetlenül  $l \rightarrow \infty$  mellett ugyanazon számhoz, a várható értékhez konvergál. Egy TAR modell több fixponttal rendelkezhet, vagyis az, hogy az  $l$ -lépéses előrejelzés mihez konvergál, függhet az idősor konkrét értékeitől. Ezen értékek olyan összességét, mely adott fixponthoz való konvergenciát eredményez, a fixpont vonzási tartományának nevezzük. Ezen kívül határciklus is felléphet, amikor az előrejelzés nem egy számhoz, hanem egy periodikusan ismétlődő számsorozathoz tart. A jelenlegi VTAR modellünk három fixponttal és egy határciklussal rendelkezik. A határciklus periódusideje 3400 év, ami a 28 DO esemény között eltelt átlagos idővel esik egybe. A megfigyelt adatsor a teljes időszakban mindössze négyszer esett a határciklus vonzási tartományába, ám éppen akkor, amikor két egymást követő DO esemény között szinte pontosan 3400 év telt el. A három fixpontot a fixpontok vonzási tartományának előfordulásai idejének függvényében ábrázolva, az adatsort jellemző trend szakaszonként konstans közelítését kapjuk.

Megjegyezzük, hogy a VTAR modell meteorológiai alkalmazásaival kapcsolatban csak az egyváltozós esetre (TAR) találtunk példát, de saját korábbi tanulmányaink (Matyasovszky, 2001; 2003) kivételével itt is csupán Zwiers and von Storch (1990) egy igen egyszerű módszere említhető.

7.

Feladatul tűztük ki a hazánkban nagyon elterjedt parlagfű erősen allergén pollenjének napi koncentráció becslését. Szeged, Legnano és Lyon napi parlagfű pollenkoncentrációit hoztuk kapcsolatba (Makra et al., 2011a) a megelőző napi koncentrációval és a megelőző napi átlaghőmérséklettel, csapadékösszeggel és átlagos szélességgel az 1997-2006 időszakban egy időfüggő nemparaméteres regressziós technika, az ún. time-varying coefficient model (Cai, 2007) segítségével. Itt ismét megemlíthetjük, hogy az eljárás korábbi meteorológiai alkalmazásáról nincsen tudomásunk. A Szegeden kívüli további két város bevonására azért került sor, mert a Kárpát-medencén kívül még a Pó-alföld (Legnano) és a Rajna völgye (Lyon) Európa erősen parlagfüves területei (Makra et al., 2011a).

A becslés által megmagyarázott relatív variancia Szegedre a legnagyobb (52,2%) és Legnanora a legkisebb (22%), tehát a legpontosabban Szeged napi parlagfű pollenkoncentrációja becsülhető a három hely közül. A legfontosabb meteorológiai változónak a napi középhőmérséklet (Szeged és Legnano) és a napi csapadék (Lyon) bizonyult.

8.

Szegedre a parlagfű napi koncentráció becslését időfüggő medián és kvantilis regresszió (Koenker and Bassett, 1978) segítségével is elvégeztük (Makra and Matyasovszky, 2010). A medián regresszió akkor igazán hasznos, amikor a becslendő változó valószínűségi eloszlása erősen aszimmetrikus és így a medián és a várható érték jelentősen különbözik. A gyakorlati feladatok során ugyanis nem annyira a minél kisebb négyzetes hiba, hanem a minél alacsonyabb abszolút hiba biztosítása a cél, és a medián regresszió éppen az abszolút hibaátlagot minimalizálja. A kvantilis regressziót számos valószínűségi érték mellett végrehajtva képet kapunk a prediktandusznak (napi parlagfű koncentráció) a prediktorok

(előző napi koncentráció és meteorológiai változók) melletti feltételes valószínűségi eloszlásáról is.

A medián regresszió természetesen számottevően (20,9%-kal) kisebb abszolút hibaátlagot nyújtott az előző pontban említett regresszióhoz képest (a megmagyarázott variancia csökkenése árán). A kvantilis regressziót ezúttal az esős és száraz napokra szétválasztva külön-külön értelmeztük a csapadék szerepének egyszerűbb kezelése érdekében. A napi parlagfű pollenkoncentráció kvantilisei általában kisebbek az esős, mint a száraz napokon, továbbá a napi koncentráció valószínűségi eloszlása sokkal elnyújtottabb a magas koncentrációk felé a száraz napokon. Az esős napokhoz tartozó kvantilisek azt jelzik, hogy a pollenkoncentrációk jóval kisebb változékonyságúak a csapadékos napokon. Mindez világosan jelzi a csapadék koncentrációcsökkentő hatását.

A hazai parlagfű pollenterhelés súlyosságára jól rávilágít a kvantilis regressziónak a nulla valószínűségi érték melletti alkalmazása. Ezzel tulajdonképpen a koncentrációk alsó határát lehet meghatározni, mert ez a kvantilis az a legnagyobb koncentráció, amelynél nagyobb koncentráció egy valószínűséggel fordul elő. Az egészségi kockázatot jelentő kritikus parlagfű koncentrációt 20 pollenszám  $\text{m}^{-3}$ -nek véve (Jäger, 1998), a lehetséges legkisebb koncentráció csaknem 20 napon át bizonyosan meghaladja az említett küszöbértéket még úgy is, hogy ha pusztán az évi menettel foglalkozunk, vagyis a prediktorok értékét, tehát például egy előző napi esetlegesen magas koncentrációt, figyelembe sem vesszük.

9.

Az előző vizsgálatok során kiderült, hogy a parlagfű pollenkoncentráció becsléséhez a leghasznosabb prediktor a megelőző nap koncentrációja. Célszerű ezért egy AR(1) folyamatot illeszteni a koncentrációk idősorához. Egy  $p$ -edrendű stacionárius autoregresszív (AR( $p$ ))

folyamat a stacionárius folyamatok igen tág körét tetszőleges pontossággal közelíti megfelelő  $p$  mellett abban az értelemben, hogy az AR modell által generált kovarianciafüggvény tetszőlegesen közel van a modellezni kívánt valóságos folyamat kovarianciafüggvényéhez. Ha a szóban forgó idősor Gauss-folyamatból származik, akkor az AR( $p$ ) modell minden statisztikai jellemző (és nem csak a kovarianciafüggvény) szempontjából jó közelítése az idősort generáló folyamatnak. Ha azonban a modellezendő folyamat nem gaussi, akkor ez az utóbbi megállapítás általában nem érvényes. Mivel a napi pollenkoncentrációk bizonyosan nem normális eloszlásúak, a hagyományos AR modellezés nem igazán célra vezető.

Comtois (2000) munkájával szemben kimutattuk, hogy a napi parlagfű pollenkoncentrációk jó közelítéssel időben változó paraméterű lognormális eloszlásúnak tekinthetők (Matyasovszky and Makra, 2011). Ennek kihasználásával kidolgoztunk egy időfüggő lognormális AR(1) modellt, amely az aktuális nap koncentrációjának az előző napi koncentrációra vonatkozó időfüggő feltételes valószínűségi eloszlását adja meg. Előállítottuk a modell egy olyan kiterjesztését is, amelyben az említett eloszlás az előző napi meteorológiai változók (most csak a napi középhőmérséklet) értékeitől is függ. A modell által megmagyarázott variancia 53,5%, míg a megmagyarázott varianciához hasonlóan értelmezett megmagyarázott abszolút hibaátlag 40,3%. A 7. pontban említett módszerrel nyert hasonló értékek 52,2%, illetve 37,4%. Ezek a számok ugyan azt sugallják, hogy a mostani eljárás nem hoz gyökeres javulást az előzőekhez képest, de valójában lényegesen gazdagabb információt szolgáltat. Nevezetesen, a kiterjesztett modell által nyújtott feltételes eloszlás nem csupán pontbecslést (lásd 7. pont), hanem intervallumbecslést is lehetővé tesz. Könnyen kiszámolható ugyanis adott napi pollenkoncentráció és napi középhőmérséklet esetében a másnapi koncentráció adott intervallumba esésének becsült valószínűsége. A legérdekesebb kérdés talán az, hogy milyen valószínűséggel számíthatunk valamilyen kritikus koncentráció meghaladására. Az értekezésben bemutatott ilyen eredményeink igen meggyőzőknek tűnnek.

10.

Az elmúlt néhány évized során számos növényi faj pollenje által kiváltott allergiás tünetek és allergiás légúti betegségek számának erőteljes növekedése figyelhető meg világszerte (Damialis et al., 2007). Az ezzel párhuzamosan zajló globális éghajlatváltozás miatt logikus felvetés, hogy a pollenszezon fenológiai jellemzői (a pollenszezon kezdete, vége, tartama) és mennyiségi jellemzői (évi összes pollenszám, éves napi csúcspollen) is változást mutatnak.

A kérdést vizsgálva 19 taxon napi pollenszámait elemeztük a rendelkezésünkre álló 1997-2007 közötti 11 éves időszakban Szegedre (Makra et al., 2011b). Mivel a vizsgált pollen karakterisztikák valószínűleg nem tekinthetők normális eloszlásúaknak (például az éves csúcspollen), ezért a jó ismert *t*-próba helyett a Mann-Kendall-tesztet (MK-tesztet) (Önöz and Bayazit, 2003) alkalmaztuk. A trend létének igazolása mindössze 11 adat felhasználásával persze igen kevésbé ígérkezett sikeresnek. Valóban, a 19 taxon 5 karakterisztikájára elvégzett összesen 95 vizsgálat mindössze 16, 10 és 3 esetben jelzett trendet a 10, 5 és 1 %-os szignifikancia-szinten, illetve pusztán az évi összes pollenszám esetében a 19 taxonra csupán 4, 1 és 0 esetben mutatkozott szignifikáns trend az említett szinteken. Ezért a napi pollenszámokra a 11 éves időszak 11 adata alapján az év összes napjára (a pollenszezonra) külön-külön elvégeztük az MK-tesztet. A trend létezéséről szóló döntés most azonos azzal a problémával, hogy a napi MK-teszt értékek évi átlaga szignifikánsan különbözik-e nullától. Ekkor az 5%-os valószínűségi szinten már 11 taxon évi összes pollenszáma mutatott szignifikáns trendet, s e 11-ből 7 növekedést jelzett. Megtörténhet azonban, hogy a pollenszezon pozitív és negatív trendeket mutató időszakokból áll össze és az MK-teszt értékek átlaga emiatt nem ad átfogó (teljes évre számított) trendet 5, 8 és 11 taxonra a 10, 5 és 1 %-os szinten. Ezt a lehetőséget azzal vizsgáltuk, hogy a napi MK-teszt értékek évi menetét nemparaméteres módszerrel becsültük, ami már az összes taxonra trendet jelezett. A nemparaméteres regresszió a napi bontású MK-statisztikákon keresztül tehát

összehasonlíthatatlanul finomabb képet nyújt az allergén pollenek trendjéről, mint a mások által követett szokványos eljárás.

Meteorológiai változók (minimum hőmérséklet, maximum hőmérséklet, középhőmérséklet, globálsugárzás, relatív nedvesség, szélsébség és a csapadékösszeg) napi értékeire hasonló vizsgálatot végeztünk, s ezek 11 éves trendjének évi menetét kapcsolatba hoztuk az egyes taxonok 11 éves trendjének évi menetével. Azt találtuk, hogy az egyes meteorológiai változók trendjei, figyelembe véve a taxonok klimatikus igényeit, igen jól magyarázzák a pollenkoncentrációk trendjeit. A taxonok trendjének évi menete és a meteorológiai változók trendjeinek évi menete közötti többszörös korreláció meglepően nagyra adódott. A legnagyobb többszörös korreláció az *Artemesia* esetén 0,998, de az *Ambrosia* és *Urtica* esetén fellépő legalacsonyabb 0,827-es érték is igen magas.

## 4. Irodalom

(idézett saját közlemények vastagon szedve)

- Álvarez EE, Yohai VJ, 2011: M-estimators for Isotonic Regression. arXiv: 1105.5065v1 stat.ME.
- Bradley RS, Hughes MK, Diaz HF, 2003: Climate in medieval time. *Science*, 302: 404-405.
- Cai Z, 2007: Trending time-varying coefficient time series models with serially correlated errors. *Journal of Econometrics*, 136: 163-188.
- Cline DBH, Eubank RL, Speckman PL, 1995: Nonparametric estimation of regression curves with discontinuous derivatives. *Journal of Statistical Research*, 29: 17–30.
- Comtois P, 2000: The gamma distribution as the true aerobiological probability density function (PDF). *Aerobiologia*, 16: 171-176.
- Damialis A, Halley JM, Gioulekas D, Vokou D, 2007: Long-term trends in atmospheric pollen levels in the city of Thessaloniki, Greece. *Atmospheric Environment*, 41: 7011-7021.
- Dansgaard W, Johnsen SJ, Clausen HB, Dahl-Jensen D, Gundestrup NS, Hammer CU, Hvidberg CS, Steffensen JP, Sveinbjörnsdóttir AE, Jouzel J, Bond G, 1993: Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature*, 364: 218-220.
- Engle R, 1982: Autoregressive Conditional Heteroscedasticity with Estimation of United Kingdom Inflation. *Econometrics*, 50: 987-1008.
- Glosten LR, Jagannathan R, Runkle DE, 1993: On the Relation between the Expected Value and the Volatility of the Nominal Excess Returns on Stocks. *Journal of Finance*, 48: 1779-1801.
- Granger CWJ, 1969: Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37: 424-438.
- Grove JM, 1988: *The Little Ice Age*. New York: Methuen.
- Jäger S, 1998: Global aspect of ragweed in Europe. In: *Satellite Symposium Proceedings: Ragweed in Europe*. Edited by Spieksma FThM. Proceedings of the 6<sup>th</sup> International Congress on Aerobiology. Alk-Abelló, Perugia, Italy, pp. 6-10.
- Janas D, von Sachs R, 1995: Consistency for non-linear functions of the periodogram of tapered data. *Journal of Time Series Analysis*, 16: 585-606.
- Koenker R, Bassett GB, 1978: Regression quantiles. *Econometrica*, 46: 33-50.
- Lamb HH, 1977: *Climate History and the Future, Vol. 2, Climate: Present, Past and Future*. New York: Methuen.
- Lomb NR, 1976: Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science*, 39: 447-462.
- Mann ME, Lee JM, 1996: Robust estimation of background noise and signal detection in climatic time series. *Climatic Change*, 33: 409-445.
- Makra L, ***Matyasovszky I***, 2010: Assessment of daily ragweed pollen concentration with previous-day meteorological variables using regression and quantile regression analysis for Szeged, Hungary. *Aerobiologia*, 27: 247-259.
- Makra L, ***Matyasovszky I***, Thibaudon M, Bonini M, 2011a: Forecasting ragweed pollen characteristics with nonparametric regression methods over the most polluted areas of Europe. *International Journal of Biometeorology*, 55: 361-371.
- Makra L, ***Matyasovszky I***, Deák JÁ, 2011b: Trends in the characteristics of allergenic pollen circulation in Central Europe based on the example of Szeged, Hungary. *Atmospheric Environment*, 45: 6010-6018.

- Matyasovszky I**, 1992: Nonparametric Regression Methods for Trend Estimation of Climatological Time Series. 12<sup>th</sup> International Conference on Probability and Statistics, Toronto, Canada, 1992, pp. 5-10.
- Matyasovszky I**, 1998: Non-parametric estimation of climate trends. *Időjárás*, 102: 149-158.
- Matyasovszky I**, 2001: A nonlinear approach to modeling climatological time series. *Theoretical and Applied Climatology*, 69: 139-148.
- Matyasovszky I**, 2003: The relationship between NAO and temperature in Hungary and its nonlinear connection with ENSO. *Theoretical and Applied Climatology*, 74: 69-75.
- Matyasovszky I**, 2010a: Improving the methodology for spectral analysis of climatic time series. *Theoretical and Applied Climatology*, 101: 281-287.
- Matyasovszky I**, 2010b: Milankovitch forcing in paleoclimate data. *Climate Research*, 41: 151-156.
- Matyasovszky I**, 2010c: Trends, time-dependent and nonlinear time series models for NGRIP and VOSTOK paleoclimate data. *Theoretical and Applied Climatology*, 101: 433-443.
- Matyasovszky I**, 2011: Detecting abrupt climate changes on different time scales. *Theoretical and Applied Climatology*, 105: 445-454.
- Matyasovszky I**, 2013a: Spectral analysis of unevenly spaced climatological time series. *Theoretical and Applied Climatology*, 111: 371-378.
- Matyasovszky I**, 2013b: Estimating red noise spectra of climatological time series. *Időjárás*, accepted.
- Matyasovszky I**, Makra L, 2011: Autoregressive modelling of daily ragweed pollen concentrations for Szeged in Hungary. *Theoretical and Applied Climatology*, 104: 277-283.
- Matyasovszky I**, Ljungqvist FC, 2012: Abrupt temperature changes during the last 1,500 years. *Theoretical and Applied Climatology*, Doi: 10.1007/s00704-012-0725-8.
- Mudelsee M, 2002: TAUEST: a computer program for estimating persistence in unevenly spaced weather/climate time series. *Computers & Geosciences*, 28: 69-72.
- Nicolay S, Mabile G, Fettweis X, Erpicum M, 2008: 30 and 43 months period cycles found in air temperature time series using the Morlet wavelet method. *Clymate Dynamics*, Doi: 10.1007/s00382-008-0484-5.
- Önöz B, Bayazit M, 2003: The power of statistical tests for trend detection. *Turkish Journal of Engineering and Environmental Sciences*, 27: 247-251.
- Scargle JD, 1982: Studies in astronomical time series analysis II: statistical aspects of spectral analysis of unevenly spaced data. *The Astrophysical Journal*, 261: 835-853.
- Schulz M, Mudelsee M, 2002: REDFIT: estimating red-noise spectra directly from unevenly spaced paleoclimatic time series. *Computers & Geosciences*, 28: 421-426.
- Seidov D, Barron E, Haupt BJ, 2001: Meltwater and global ocean conveyor: northern versus southern connections. *Global and Planetary Change*, 30: 257-270.
- Severinghaus JP, 2009: Climate change: southern see-saw seen. *Nature*, 457, Doi: 10.1038/4571093a.
- Steig EJ, Alley RB, 2002: Phase relationships between Antarctic and Greenland climate records. *Annals of Glaciology*, 35: 451-456.
- Tsay RF, 1998: Testing and Modeling Multivariate Threshold Models. *Journal of the American Statistical Association*, 93: 1188-1202.
- Wishart J, 2009: Kink estimation with correlated noise. *Journal of Korean Statistical Society*, 38: 131-143.
- Zhao O, Woodroffe M, 2012: Estimating a monotone trend. *Statistica Sinica*, 22: 359-378.
- Zwiers F, von Storch H, 1990: Regime dependent auto-regressive time series modeling of the Southern Oscillation. *Journal of Climate*, 3: 1347-1360.



**5. Az értekezés eredményeihez kapcsolódó nem idézett saját közlemények**

- Deák JÁ, Makra L, **Matyasovszky I**, Csépe Z, Muladi B, 2013: Climate sensitivity of allergenic taxa in Central Europe associated with new climate change – related forces. *Science of the Total Environment*, 442: 36-47.
- Matyasovszky I**, Makra L, Csépe Z, 2012: Potential reasons of day-to-day variations of ragweed (*Ambrosia*) pollen counts in association with meteorological elements. *Arhiv Za Higijenu Rada I Toksikologiju*, 63: 311-320.
- Matyasovszky I**, Makra L, 2012: Estimating extreme daily pollen loads for Szeged, Hungary using previous-day meteorological variables. *Aerobiologia*, 28: 337-346.
- Makra L, **Matyasovszky I**, Páldy A, Deák AJ, 2012: The influence of extreme high and low temperatures and precipitation totals on pollen seasons of *Ambrosia*, *Poaceae* and *Populus* in Szeged, southern Hungary. *Grana*, 51: 215-227.
- Matyasovszky I**, Makra L, Guba Z, Pátkai Zs, Páldy A, Sümeghy Z, 2011: Estimating the daily *Poaceae* (grass) pollen concentration in Hungary by linear regression conditioning on weather types. *Grana*, 50: 208-216.
- Makra L, Sánta T, **Matyasovszky I**, Damialis A, Karatzas K, Bergmanns K-C, Vokou D, 2010: Airborne pollen in three European cities: Detection of atmospheric pathways by applying 3D clustering of backward trajectories. *Journal of Geophysical Research*, 115, D24220, 16 p, doi:10.1029/2010JD014743.