

Óbudai Egyetem

Doktori (PhD) értekezés
tézisfüzete



Adatbányászati és térinformatikai módszerek
biológiai és meteorológiai alkalmazásokkal

Gimesi László

Témavezetők:

Prof. Dr. Galántai Aurél egyetemi tanár

Prof. Dr. Szeidl László egyetemi tanár

Alkalmazott Informatikai Doktori Iskola

Budapest, 2011.

1. A kutatás előzményei

Napjainkra az informatika és alkalmazásainak fejlődése lehetővé tette, hogy óriási mennyiségű adatot halmozzunk fel különböző adatbázisokban. A világ adatmennyisége exponenciálisan nő, évente nagyjából megduplázódik (Adriaans & Zantinge 2002). Megfelelő eszközök (hardver, szoftver, modellek, algoritmusok, stb.) nélkül megoldhatatlan lenne a nagy mennyiségű adat feldolgozása, a számunkra fontos információ kiválasztása, megőrzése, illetve a szükségtelen adatok kiszűrése (Detrekői & Szabó 2002). Az adatok gyűjtéséhez és tárolásához már ma is rendelkezünk alkalmas eszközökkel, de az elemzésükre és felhasználásukra még kevés a jól használható módszer (Compieta et al. 2007).

Az informatikai kutatások egyik céljává vált, hogy az egyre duzzadó adathalmazokból a lehető legtöbb hasznos információt lehessen kinyerni (Adriaans & Zantinge 2002).

Az adatbányászat segítségével az adatbázisokból – a gyakorlatban is jól hasznosítható – rejtett információkat, összefüggéseket, szabályszerűségeket, mintázatokat nyerhetünk ki (Fajszi et al. 2010). Azért, hogy az adatbányászat folyamatát megkönnyítsük, első lépésként kiválasztjuk a kutatásunkhoz szükséges releváns adatokat (Adriaans & Zantige 2002), vagyis adattárházat (adatpiacot) hozunk létre (Han & Kambee 2004). Ennek során az adatbázisok egyesítését és szűrését is elvégezzük (Bogdanova & Georgieva 2008).

Az adattárház összeállítása után lehetővé válik, hogy az adatok között keressünk, a kapott információkat különböző szempontok szerint kiértékeljük, azokon elemzéseket végezzünk. Az adatok elemzésének legelterjedtebb módszere a leíró statisztika, de egyre gyakrabban alkalmazzák a klaszteranalízist, a regresszió-elemzést és a mesterséges neurális hálókat is (Fajszi & Cser 2004, Han & Kambee 2004).

Az emberi gondolkodáshoz közelebb áll a képi megjelenítés, mint a nagy, numerikus adatokat tartalmazó táblázatok, amelyek ugyan pontos információt adnak, de nehezen kezelhetők, és az összefüggések bemutatására sem alkalmasak. Ezért fontos, hogy az adatbányászat során kapott eredményeket minél szemléletesebben jelenítsük meg. Ennek megfelelően az adatbányászat erősen támaszkodik a vizualizációra, mint munkaeszközre (Fajszi et al. 2010).

Az adatbányászat egyik speciális területének tekinthető a térinformatika, ahol földrajzi helyhez kötött információkat dolgozunk fel (Maguire 1991). A helyhez kötött, térbeli információk feldolgozására használt rendszereket térinformációs rendszereknek nevezzük, ezek elméletével, feldolgozásával foglalkozó tudomány a térinformatika (GIS – Geographical Information System).

Ahogy az adatbányászatban, úgy a térinformatikában is fontos szerepe van az időnek (Peuquet 1999), hiszen ugyanarról a helyről, ugyanarról az objektumról többször is gyűjthetünk adatokat, így azok történetét, időbeli változását is nyomon követhetjük. Azokat a rendszereket, ahol az időbeli változásokat követik, monitoring rendszereknek nevezzük (Detrekői & Szabó 2002).

Az értekezésben ismertetett adatbányászati módszerek bemutatásához olyan adatokat kerestünk, amelyek még kevésbé feltártak, illetve hosszú idősorokat tartalmaznak. Az adatok kiválasztásában döntő szerepet játszott az is, hogy napjainkban valamennyi tudományterületen központi kérdéssé vált a klímaváltozás és annak vizsgálata. A kutatás jelentőségét az adja, hogy a klímaváltozás hatással van a különböző gazdasági, ökológiai, mezőgazdasági, hidrológiai stb. folyamatokra és ezeken keresztül az emberi élet minőségére és fenntarthatóságára.

A feladat jellege, természete, komplexitása szükségessé tette, hogy a legkorszerűbb technikákat dolgozzuk ki, illetve használjuk. Ezen informatikai és matematikai eszközök nélkül a feladatot nem tudtuk volna megoldani.

A klímaváltozás kutatásához a szakirodalmak elsősorban az időjárási (csapadék, hőmérséklet, stb.) adatokat használják, és ezek segítségével szemléltetik annak irányát és mértékét. Ezen kívül egyre gyakrabban használnak biológiai indikátorokat, ahol a populációdinamikai idősorokból következtetnek a környezetben bekövetkező változásokra.

Az értekezésben meteorológiai adatforrásként Ferenczy (2008) által közzétett „Kutdiak” adatbázist, Szenteleki (2007) valamint Szenteleki és munkatársai (2007) által publikált „KKT” adatbázist, illetve az Országos Meteorológiai Szolgálat (2008) honlapján közzétett budapesti, debreceni és szegedi adatokat használtuk fel.

Biológiai indikátorként az Országos Növényvédelmi és Erdészeti fénycsapda-hálózat rovaradatai közül a *Lepidoptera* rendet választottuk. Ezt elterjedésük, fajgazdagságuk, jelentős egyedszámuk és a rendelkezésre álló jelentős mennyiségű és több éves megfigyelésből származó adatsoruk indokolta. Ezen kívül a lepkék jó indikátorok, mivel a környezeti hatások megváltozására érzékenyen reagálnak (Hufnagel et al. 2008).

A biológiai adatok elemzéséhez, valamint a meteorológiai adatokkal való összefüggések feltárásához a szakirodalmakban alkalmazott statisztikai módszereket használtunk, illetve továbbfejlesztettük azokat.

A **diverzitási-indexek** (diverzitás mérőszámai) a fajgyakoriságok vagy faj előfordulási valószínűségek halmazára értelmezett számszerű függvények (Izsák 2001). Tehát egy életközösség diverzitása – ökológiai értelemben – a fajok számának és abundanciájának valamilyen függvénye. A diverzitási-indexek azonban nem adnak információt az egyedek térbeli elhelyez-

kedéséről, ami legalább annyira jellemezheti a populációt, mint a fajok száma, vagy a diverzitás (Menhinick 1962). Diverzitási mérőszámként nagyszámú függvényt alkalmaznak a statisztikus ökológiában (Dewar & Porté 2008, Izsák 2001, Mishra et al. 2009, Sipkay et al. 2005, Tóthmérész 1997). A publikációkban leírt különböző diverzitási-indexek egy adott fajközösség diverzitását más és más oldalról mutatják be. Általános tapasztalat, hogy nagyszámú állat- vagy növényközösség különböző indexszel mért diverzitása jelentős pozitív korrelációt mutat. Ennek ellenére a többféle diverzitási-index alkalmazását éppen az indokolhatja, hogy azok a diverzitási viszonyokat más-más oldalról világítják meg (Izsák 2005).

A nagy szakirodalommal rendelkező fajdiverzitás vizsgálatok mellett felmerült az igény a **fajabundancia-modellek** kidolgozására (Izsák & Szeidl 2009), azaz annak vizsgálatára, hogy a fajok milyen egyedszámban (tömegességben) vesznek részt az adott közösségben (Bartha et al. 2007, Magurran 1988). A fajabundancia-modellek esetében a fajok közötti erőviszonyokat ábrázoljuk, ami szintén jellemezheti a klímaváltozást, hiszen a dominancia változása utalhat a környezeti változásokra.

Az ismertetett két modell mellett használják a **kvantilis regressziót** is, amely hatékony eszköz a változások szemléltetésére (Cade et al. 2005, Helmus et al. 2007). A lineáris kvantilis regresszió egyszerűen szemlélteti a bekövetkező változások arányát, és segítségével olyan összefüggések is felfedezhetők, amelyek a tradicionális statisztikai módszerekkel nem (Chamailé-Jammes et al. 2007).

Az értekezésben egy olyan módszertani, informatikai rendszert mutatunk be és fejlesztünk tovább, amely segítségével az ökológiai folyamatok modellezhetők, vizsgálhatók és megjeleníthetők. Az ismertetett módszerek segítségével az adatokat elemeztük, valamint összefüggéseket kerestünk a klímaadatok (elsősorban a hőmérséklet) és különböző közösségek dinamikája között.

2. Célkitűzések

Az értekezés célja az adatbányászati és térinformatikai eszközök felhasználásával meteorológiai és biológiai hosszú idősorok vizsgálata, valamint olyan módszertani, informatikai rendszer kidolgozása, bemutatása, amely segítségével az ökológiai folyamatok vizsgálhatók és megjeleníthetők. Fontos, hogy a lehető legpontosabban tudjuk meghatározni és legszemléletesebben bemutatni a közösségekben bekövetkező változásokat és azok tendenciáit. Ehhez a lehető leghosszabb és a lehető legnagyobb földrajzi területet lefedő adatsorra van szükségünk.

Célunk volt Magyarország különböző helyeiről gyűjtött, különböző formában rendelkezésre álló meteorológiai és rovaradatokról egy adattárház létrehozása, amely segítségével lehetővé válik az adatok különböző szempontok szerinti elemzése. Az adattárház összeállításánál törekedtünk arra, hogy biztosítsuk a lehető legegyszerűbb és a lehető legszélesebb körű felhasználhatóságot.

A számítási eredmények elemzésében segítséget jelent, ha adatainkat könnyebben értelmezhető formában, grafikusán jelenítjük meg. Ennek érdekében egy, a térinformatikából átvett, háromdimenziós módszer kidolgozásával elértük, hogy az adatokat olyan idősor-ábrán tudjuk bemutatni, ahol együtt vizsgálható a hosszútávú és a szezonális változás.

Az adatbázis létrehozása és a megjelenítő módszer kidolgozása után válik lehetővé, hogy adatainkat különböző statisztikai módszerrel elemezzük, és az eredményeinket könnyen értelmezhető formában megjelenítsük.

Az értekezésben biológiai idősorokat vizsgáltunk különböző adatbányászati és térinformatikai eszközök segítségével, valamint módszertani rendszert javasoltunk az ökológiai folyamatok vizsgálatára.

Az adatelemzés elsődleges célja a közösségszinten megmutatkozó sokéves tendenciák, a szezonális közösségdinamikai mintázatok, valamint ezek hőmérsékletfüggésének elemzése, a szakirodalomokban fellelhető fajabundancia modellek lepkékre való alkalmazhatóságának igazolása, illetve a klímaváltozás hatásainak bemutatása a populáció viselkedésére.

Az értekezésben kidolgozott témával jelentős hazai és nemzetközi tudományos iskolák foglalkoznak, és produkálnak újabb és újabb eredményeket. Egy interdiszciplináris (informatika, statisztika, biológia) tudományról van szó, ezért a felvetett problémák igen bonyolultak, a megoldásukhoz szükséges tudás szerteágazó, és sok nézőpontot igényel. Az informatika fejlődése számos új lehetőséget nyújt, egyre összetettebb módszerek megoldására ad lehetőséget.

Az értekezésben kitűzött célunkat a korszerű informatikai eszközök nélkül nem tudtuk volna megvalósítani.

3. Vizsgálati módszerek

Az Országos Növényvédelmi és Erdészeti fénycsapda-hálózat adatait dBase formában kaptuk meg. Feladatunk volt, hogy az ebben a formában nehezen feldolgozható adatokból egy olyan, a további kutatásokhoz nélkülözhetetlen adatbázist hozzunk létre, amely lehetővé teszi az egyszerű, általános formátumú adatelérést. Ennek érdekében PC-s számítógépes környezetet és az MS Office programhátteret választottuk, ezért az adatbázist Access alapon készítettük el, amelynek az az előnye, hogy biztosítja az SQL adatelérést, valamint egyszerű konvertálhatóságot tesz lehetővé. Így megteremtettük a lehető legszélesebb körű felhasználás lehetőségét.

A meteorológiai adatok esetében a kapott adatok típusa sem volt egységes. Azokhoz Access, Excel illetve szöveges formában jutottunk hozzá.

A kapott adatokat az adatbányászat módszereit alkalmazva dolgoztuk fel. Első lépésként a rendelkezésre álló adatbázisokból adattárházat hoztunk létre. Ez a folyamat magába foglalja az adatbázisok egyesítését és szűrését. Ehhez Visual Basic nyelvű programokat fejlesztettünk ki. E programozási nyelvet azért választottuk, mert így lehetőségünk volt a közvetlen szöveges, Excel és Access állományok elérésére.

Az adatbázisok egyesítésével egy olyan 4 dimenziós adatstruktúrát hoztunk létre, amely biztosítja a csapdakód, a lepkefaj-kód, az évenkénti és a naponkénti keresést, kigyűjtést. Az adattárház létrehozásával egy időben az automatikusan elvégezhető szűréseket is végrehajtottuk, amelynek során eltávolítottuk az adatrögzítési hibából eredő rossz dátumú, faj- és csapdakódú, illetve a duplán rögzített adatokat.

A populációdinamikai vizsgálatok céljára – a szakirodalmak alapján – kiválasztottuk azokat a csapdákat, amelyek azonos helyen, hosszú ideje üzemeltek megszakítás nélkül. E csapdák adatainak egyesítésére új módszert dolgoztunk ki, amelyhez a mozgóátlag módszert használtuk fel.

Az adattisztítás és a szűrések befejezése után a lepkeadatbázis 9 csapda, 281 faj, 44 éves napi adatait tartalmazza, ami összesen 4.020.614 rekordot jelent.

Meteorológiai adatforrásként – a bevezetőben említett – Ferenczy (2008), Szenteleki (2007) valamint Szenteleki és munkatársai (2007) által közzétett adatbázisokat használtuk. Az adatbázisok szűréséhez, ellenőrzéséhez felhasználtuk az Országos Meteorológiai Szolgálat (2008) honlapján közzétett budapesti, debreceni és szegedi adatokat. Az adatbázis összeállításánál a vizsgált csapdákhöz legközelebbi meteorológiai állomás adatait vettük figyelembe. Az állomásokon mért napi középhőmérsékletből számoltuk az országos átlag-középhőmérsékletet,

amely meghatározásánál – ahogy a csapdaadatok esetében is tettük –, 9 állomás adatából, 9 napos mozgóátlagot számoltunk.

A számítási eredmények elemzésében segítséget jelent, ha adatainkat könnyebben értelmezhető formában, grafikusán jelenítjük meg. Ennek érdekében a térinformatikából (GIS) átvett háromdimenziós felületmodellt használtuk. E módszer segítségével az eredményeket olyan idősor-ábrán tudjuk bemutatni, ahol együtt vizsgálható az éves és az éven belüli (szezonális) változás.

A hosszú idősorok ábrázolására használt háromdimenziós vizualizációs módszer része a felület simítása, amelyhez neurális hálózatot, valamint különböző interpolációs eljárásokat alkalmaztunk, illetve fejlesztettünk ki. A legkisebb négyzetek módszerével történő közelítésre saját eljárást és Visual Basic programot dolgoztunk ki. A programmal kapott numerikus adatok megjelenítésére az AutoCAD (Autodesk Inc.) és az ArcGIS (GIS by ESRI) programokat is használtuk. A háromdimenziós megjelenítéshez az AutoCAD-hez LISP nyelvű programot fejlesztettünk ki.

A fajoknak egy adott területre és időszakra vonatkozó változatossága biológiai diverzitással jellemezhető, amely vizsgálatára – a szakirodalmak alapján – számos módszert dolgoztak ki. Mi elsősorban a fajszám, az egyedszám és a Shannon-féle diverzitási-indexet használtunk, de bemutatunk és összehasonlítottunk más diverzitási-indexeket is.

A diverzitási-indexek kiszámításához Hammer és munkatársai (2001) által publikált PAST (Paleontological Statistics) programmal dolgoztunk. Az adatok leválogatásához Access és Excel makrókat készítettünk. A programokat különböző diverzitási-indexekre futtattuk, és a kapott eredményeket az általunk kidolgozott idősor-ábrákon mutatjuk be.

A fajabundancia-modellek abból a feltevésből indulnak ki, hogy ha egy faj (közösség) nagyobb forráshoz (tápanyag, élettér stb.) jut, akkor nagyobb egyedszámban is jelenik meg (Izsák & Szeidl 2009). Az ökológiával foglalkozó szakirodalmak három alapmodellt: geometriai, lognormális és szimultán pálcatorési modelleket említenek, de egyes publikációkban más modellek is fellelhetők.

Az értekezésben különböző modelleket hasonlítottunk össze a tényadatokkal. Az elemzést többéves és havi bontásban is elvégeztük, amihez Excel makrókat készítettünk. A függvények összehasonlítását a legkisebb négyzetek módszerével végeztük. A hagyományos statisztikai elemzésekhez és a grafikonok megrajzolásához az Excel adatelemző és diagramkészítő modulját használtuk.

A hosszú távú változások bemutatására a kvantilis regressziót alkalmaztuk, amely a mérési adatok adott hányadának (például: 10, 20, 30%-os) változását szemlélteti.

Az elemzéshez Excel makrót készítettünk, valamint felhasználtuk a GRETL (GNU Regression, Econometric and Time-series Library) programot. A grafikonok megjelenítéséhez az Excel programmal dolgoztunk.

A szezonális dinamika vizsgálatához a már említett diverzitási-indexeket és fajabundancia-modelleket használtunk. Az eredményeket gyakoriság vizsgálattal és különböző osztályozási módszerek segítségével is alátámasztottuk. Ennek érdekében elkészítettük a havi befogások dendrogramját (clusterét) és a nem-metrikus többdimenziós skálázást (NMDS) is.

A fajok hőmérsékletfüggését hagyományos statisztikai módszerekkel vizsgáltuk.

A fajok területi elosztásának bemutatásához az AutoCAD és az ArcGIS programokat használtuk fel.

4. Új tudományos eredmények

I. Téziscsoport: Adatbázis (adattárház) létrehozása, a vizsgálathoz kidolgozott adatbázis eljárási eljárások és informatikai eszközök

- I.1. Magyarország különböző helyeiről gyűjtött, különböző formában rendelkezésre álló rovar- és meteorológiai adatokból adattárházat hoztam létre, amely biztosítja a legelterjedtebb adatkezelő programok (Excel, Access, SQL) számára az adatelérést. Az adatok egyesítéséhez, szűréséhez Visual Basic nyelvű programokat fejlesztettem ki, melyek lehetőséget nyújtanak közvetlen szöveges, Excel és Access állományok megfelelő mélységű elérésére. Kidolgoztam egy egyesítési módszert, amely a különböző helyeken mért adatokat egységes, országos adatsorrá konvertálja. [6]
- I.2. A hosszú idősorok ábrázolására, az éven belüli és a hosszú távú tendenciák vizsgálatára háromdimenziós vizualizációs módszert dolgoztam ki. A felület simításához különböző interpolációs eljárásokat, valamint neurális hálózatot használtam. Saját eljárást és Visual Basic programot dolgoztam ki a legkisebb négyzetek módszerével történő közelítésre. A programmal kapott numerikus adatok megjelenítésére az AutoCAD (Autodesk Inc.) és az ArcGIS (GIS by ESRI) programokat is felhasználtam. A háromdimenziós megjelenítéshez AutoCAD – LISP nyelvű programot fejlesztettem. A módszer és a kifejlesztett eszközök alkalmazásaként Magyarországon egyedülként, 281 lepkefaj 3D-s idősor-ábráját készítettem el. [1, 2, 3, 4, 5]
- I.3. A diverzitási indexek vizsgálatához a PAST programot használtam. Az adatok leválogatásához Access és Excel makrókat készítettem. A programokat különböző diverzitási indexekre futtattam, és a kapott eredményeket az általam kidolgozott eljárással vizualizáltam, amelynek alapján a rejtett változási tendenciák kimutathatóvá váltak.
- I.4. A fajabundancia modellek és a tényadatok elemzéséhez, összehasonlításához Excel makrót készítettem. A függvények összehasonlítását a legkisebb négyzetek módszerével vizsgáltam. A hagyományos statisztikai elemzésekhez és a grafikonok megrajzolásához az Excel adatelemző modulját használtam. A lineáris kvantilis regresszió számításához Excel makrót készítettem, a megjelenítés az Excel grafikonrajzolójával történt. A fajok területi eloszlásának bemutatásához szintén az AutoCAD és az ArcGIS programokat használtam.

II. Téziscsoport: Magyarországi rovar- és meteorológiai adatok együttes vizsgálata a létrehozott adattárház, adatbányászati és informatikai eszközök felhasználásával

II.1. Diverzitási idősorok. Az egyes diverzitási-indexeket összehasonlítva megállapítottam, hogy a különböző modelleket használva hasonló eredményt kapunk. Mindegyik modell alkalmas a közösségek szezonális és hosszú távú dinamikájának bemutatására.

A diverzitási-indexek időbeli mintázatait vizsgálva megállapítottam, hogy a vizsgált időszakban jelentős diverzitás-szerkezeti változások történtek a viszonylag stabil egyedszám és fajszám-viszonyok ellenére. A Shannon indexben és más diverzitási mutatókban is határozottan kimutatható a nyári diverzitásemelkedés, valamint egy téli, enyhe diverzitáscsökkenés 1962 és 2006 között. [6]

II.2. Fajabundancia modellek. A szakirodalomban publikált fajabundancia modellek általában jól illeszkednek a valóságos adatokhoz, amit több példa segítségével is bemutattam. A dolgozatban példák alapján igazoltam, hogy a publikációk többségében ismertetett modellek nem biztosítják a teljes átmenetet a geometriai és a pálcatorési modell között. Ezért javasoltam a bétaeloszlás-modell használatát is, amely jobban megfelel az átmenet biztosítására.

Megállapítottam, hogy nemcsak az erőviszonyoktól, hanem a közösség egyedszámától is függ az, hogy a rangsort melyik fajabundancia modell írja le a legjobban.

Igazoltam, hogy a Magyarországi éghajlati körülmények között a lepkék fejlődésében négy jól elkülöníthető időszak figyelhető meg. Ezt a megállapítást több statisztikai modellel is alátámasztottam.

II.3. Sokéves tendenciák közösségi szinten. A közösségi szinten megmutatkozó sokéves tendenciák elemzésére készített lineáris kvantilis regressziók alapján megállapítottam, hogy az 1962 és 2006 közötti időszakban az adatok az éves összes befogott egyedszám 10, 20, ...90%-át egyre hamarabb érték el, ami a szezon 15-25 napos előretolódását jelenti. Az 1974 és 2006 évek közötti homogénebb szakaszra külön is elvégezve az elemzést, azonos eredményt kaptam.

A Shannon-diverzitás maximumának 10, 20, valamint 80, 90%-át egyre később, az 50, 60%-át viszont egyre korábban érik el az adatok az 1962 és 2006 közötti időszakban, ami a téli fajok diverzitáscsökkenését támasztja alá.

II.4. Szezonális tendenciák. A szezonális közösségdinamikai mintázatok elemzése során azt az eredményt kaptam, hogy a teljes időszakot figyelembe véve négy lokális diverzitásmaximum állapítható meg, ami négy jól elkülönülő időszakot határoz meg. A tavaszi és nyári aspektus határa áprilisban, a nyári és őszi határa pedig szeptemberben található. Az egyedszám és fajsza szám adatok lefutási mintázata ezt a jelenséget önmagában nem mutatja, az inkább a hőmérséklet lefutását követi.

Megállapítottam, hogy a különböző hónapokban a napi diverzitási értékek gyakorisági eloszlásainak az alakja is jellegzetesen eltér egymástól.

II.5. Az adatok és a hőmérséklet kapcsolata. Bizonyítottam, hogy a lepkék repülési aktivitásának és ennek következtében a befogásnak alsó és felső küszöbértéke van a hőmérséklet függvényében. Több módszerrel is igazoltam, hogy a hőmérséklet növekedésével (egy határ felett) csökken a lepkék száma és a diverzitásuk is.

Az éves hőösszeg lineáris kvantilis regressziós vizsgálatával megállapítható, hogy az alacsonyabb százalékok átlépési napjai negatív tendenciát, a magasabb százalékok átlépési napjai pedig pozitívat mutatnak, ami együttesen a vegetációs időszak hosszának növekedésére utal. Ez összhangban van a lepkék egyedszám változásában kimutatott szezon előbbre tolódásával.

A lepkebefogási adatok és a napi középhőmérséklet összefüggéseit vizsgálva megállapítottam, hogy a napi középhőmérséklet értékei és a napi diverzitás értékei szorosan korrelálnak egymással, továbbá a hőmérséklet szórása és a diverzitás egy negatív korrelációt mutat.

5. Az eredmények hasznosítási lehetősége

Az értekezésben elsősorban fénycsapdával befogott *Lepidoptera* adatokkal dolgoztunk, de bemutattunk néhány más alkalmazási területet is, ahol a leírt módszerek használhatók. Ilyenek például a meteorológiai, vízi ökológiai és elméleti adatsorok.

Az értekezés elkészítéséhez a meteorológiai adatok közül csak a hőmérsékletet függést vizsgáltuk. További kutatási téma lehet más klímaadatok hatásának elemzése.

Többek között Kúti és munkatársai (2010), Nowinszky és Puskás (2003), Wolda és munkatársai (1998) felvetik a különböző csillagászati események (naptevékenység, holdfázis, stb.) hatásait a populációdinamikára. E kutatásokhoz is felhasználhatók az általunk összeállított adatbázis és a bemutatott módszerek. Annak ellenére, hogy mi ilyen jellegű vizsgálatokat nem végeztünk, egy példát mégis bemutattunk Kúti és munkatársai (2010) által feltételezett napfolttevékenység hatásáról.

A klímaváltozás egyik kutatási területe az analóg területek keresése (Horváth 2008), amire szintén felhasználhatók az adatbázisban tárolt adatok. Igaz, az ökológiai vizsgálatok esetében ez nehezebb, mint a klímaadatoknál, hiszen egy közösség összetétele, dinamikája nagyon sok környezeti hatástól függ.

További kutatásokat igényel annak kiderítése, hogy az egyes csapdák esetén jelentkező befogási anomáliákat, illetve jelentős változásokat milyen környezeti hatások okozhatták.

Az értekezés folytatásaként tervezzük egy album kiadását, amelyben az egyes fajok változásának idősorát kívánjuk bemutatni. Ehhez Magyarországon egyedülként, 281 lepkefaj idősor-ábráját készítettük el.

Jelenleg kiadás előtt áll egy, a klímaváltozás hatásaival foglalkozó könyv, amelynek több fejezete ezen értekezés kutatási eredményei alapján készült, továbbá az értekezés alapján több publikáció megjelenését tervezzük.

Az értekezésben ismertetett háromdimenziós idősor-vizualizációs módszert – az említett felhasználásokon kívül – más szakterületeken is alkalmaztuk. Például a környezetvédelemben, CO₂ koncentráció és napfolttevékenység ábrázolására, illetve Pécs környékén végzett re-kultivációs tevékenység monitoring rendszerében a szennyezőanyagok koncentrációváltozásának megjelenítésére.

Tervezzük a Pécsi Tudományegyetemen működő szuperszámítógépen, neurális hálózat segítségével, új diverzitásmodell felállítását, valamint előrejelzést adni a közösségdinamika jövőbeni alakulására.

6. Irodalmi hivatkozások listája

- Adriaans, P., Zantinge, D. (2002): Adatbányászat. Panem Könyvkiadó.
- Bartha, S., Virágh, K., Botta-Dukát, Z., Kertész, M., Oborny, B. (2007): A fajösszetétel: textúra és kotextúra, in Pásztor, E., Oborny B. Ökológia, Nemzeti tankönyvkiadó Zrt. Budapest.
- Bogdanova, G., Georgieva, T. (2008): Using error-correcting dependencies for collaborative filtering. *Data & Knowledge Engineering*, 66: 402-413.
- Cade, B. S., Noon, B. R. Flather, C. H. (2005): Quantile regression reveals hidden bias and uncertainty in habitat models. *Ecology*, 86: 786-800.
- Chamaillé-Jammes, S., Fritza, H., Murindagomoc, F. (2007): Detecting climate changes of concern in highly variable environments: Quantile regressions reveal that droughts worsen in Hwange National Park, Zimbabwe. *Journal of Arid Environments*, 71:321-326.
- Compieta, P., Di Martinoc, S., Bertolottoa, M., Ferruccic, F., Kechadia, T. (2007): Exploratory spatio-temporal data mining and visualization. *Journal of Visual Lang. & Comp.* 18: 255-279.
- Detrekői, Á., Szabó, Gy. (2002): Térinformatika. Nemzeti tankönyvkiadó, Budapest.
- Dewar, R. C., Porté, A. (2008): Statistical mechanics unifies different ecological patterns. *Journal of Theoretical Biology*, 251: 389-403.
- Fajszi, B., Cser, L. (2004): Üzleti haszon az adatok mélyén: Adatbányászat alkalmazói szemmel. Budapesti Műszaki és Gazdaságtudományi Egyetem.
- Fajszi, B., Cser, L., Fehér, T. (2010): Üzleti haszon az adatok mélyén: Adatbányászat mindennapjai. Alinea Kiadó.
- Ferenczy, A. (2008): Napi meteorológiai jelentések Excelben. www.kutdiak.kee.hu/meteo.
- Hammer, Ř., Harper, D. A. T., Ryan, P. D. (2001): PAST: Paleontological Statistics Software Package For Education And Data Analysis. *Palaeontologia Electronica* (<http://palaeo-electronica.org>), 4: 1-9.
- Han, J., Kamber, M. (2004): Adatbányászat – Konceptiók és technikák. Panem, Budapest.
- Helmus, M. R., Bland, T. J., Williams, C. K., Ives, A. R. (2007): Phylogenetic Measures of Biodiversity. *The American Naturalist* march, 169 (3): 68-83.
- Horváth, L. (2008): Földrajzi analógia alkalmazása a klímaszcenáriók vizsgálatában. In: Harnos, Zs., Csete, L. (szerk): Klímaváltozás: környezet – kockázat – társadalom. Szaktudás Kiadó Ház, Budapest: 229-266.
- Hufnagel, L., Sipkay, Cs., Drégely-Kis, Á., Farkas, E., Türei, D., Gergócs, V., Petrányi, G., Baksa, A., Gimesi, L., Eppich, B., Dede, L., Horváth, L. (2008): Klímaváltozás, biodiverzitás és közösségökológiai folyamatok kölcsönhatásai. In: Harnos, Zs., Csete, L. (ed.): Klímaváltozás: környezet – kockázat – társadalom. Szaktudás Kiadó Ház, Budapest: 229-266.
- Izsák, J. (2001): Bevezetés a biológiai diverzitás mérésének módszertanába. Scientia, Budapest.
- Izsák, J. (2005): Populációdinamikai modellezés és fajközösségi mérőszámok. *Magyar Tudomány*, 4: 400.
- Izsák, J., Szeidl, L. (2009): Fajabundancia-eloszlási modellek, Pars Kft. Nagykovácsi.

- Kúti, Zs., Hirka, A., Petrányi, G., Szabóky, Cs., Gimesi, L., Hufnagel, L., Ladányi M. (2010): A kis téliaraszoló (*Operophtera brumata* L.) aktivitásának modellezése abiotikus paraméterekkel. *Agrárinformatika / Agricultural Informatics*, 1 (1): 40-46.
- Maguire, D. J. (1991): An Overview and Definition of GIS, in: Maguire, D. J., Goodchild, M. F., Rhind, D. W. (ed.): *Geographical Information Systems: Principles and Applications*, Longman, Harlow, 1: 9-20.
- Magurran, A. E. (1988): *Ecological Diversity and its Measurement*, Chapman and Hall, London.
- Menhinick, E. F. (1962): Comparison of invertebrate populations of soil and litter of mowed grasslands in areas treated and untreated with pesticides. *Ecology*, 43: 556-561.
- Mishra, A. K., Özger, M., Singh, V. P. (2009): An entropy-based investigation into the variability of precipitation. *Journal of Hydrology*, 370: 139-154.
- Nowinszky, L., Puskás J. (2003): A naptevékenység, in: Nowinszky, L. (ed.): *A fénycsapdázás Kézikönyve*. Savaria University Press, Szombathely: 71-74.
- Országos Meteorológiai Szolgálat (2008): Éghajlati adatsorok 1901-2000. www.met.hu.
- Peuquet, D. J. (1999): Time in GIS and geographical databases, in: Longley, P. A., Goodchild, M. F., Maguire, D. J., Rhind, D. W. (ed.): *Geographical Information Systems - Principles and Technical Issues*. New York, John Wiley, 1: 91-103.
- Sipkay, Cs., Hufnagel, L., Gaál, M. (2005): Zoocoenological state of microhabitats and its seasonal dynamics in an aquatic macroinvertebrate assembly (hydrobiological case studies on Lake Balaton, №. 1). *Applied Ecology and Environmental Research*, 3 (2): 107-137.
- Szenteleki, K. (2007): A Környezet – Kockázat – Társadalom (KLIMAKKT) Klimakutatás adatbázis kezelő rendszerei. *Klíma-21 Füzetek*, 51: 89-115.
- Szenteleki, K., Ladányi, M., Szabó, E., Horváth, L., Hufnagel, L., Révész, A. (2007): The climate research database management software. EFITA Conference, Caledonian University Glasgow.
- Tóthmérész, B. (1997): *Diverzitási rendezések*. Scientia, Budapest.
- Wolda, H., O'brien, C. W., Stockwell H. P. (1998): Weevil Diversity and Seasonality in Tropical Panama as Deduced from Light-Trap Catches (*Coleoptera: Curculionoidea*), Smithsonian Institution Press, Washington, D.C.

7. A tézispontokhoz kapcsolódó tudományos közlemények

- [1] Gimesi, L. (2004): Neurális hálózatok és térinformatikai alkalmazásuk. *Térinformatika*, 7: 13-15.
- [2] Gimesi, L. (2004): Mesterséges intelligencia alkalmazása a rekultivációban. *Acta Agraria Kaposváriensis*, 8 (3): 1-9.
- [3] Gimesi, L. (2006). A csapadékváltozás tendenciái Magyarországon. *Acta Agraria Kaposváriensis*, Kaposvári Egyetem, 10 (3): 1-10.
- [4] Gimesi, L. (2008): The use of 3-dimensional graphic display method for presenting the changes in weather. *Applied Ecology And Environmental Research* 6 (1): 165-176.
- [5] Gimesi, L. (2009): Development of a visualization method suitable to present the tendencies of the changes in precipitation. *Journal of Hydrology*, 377: 185-190.
- [6] Gimesi, L., Hufnagel, L (2010): The possibilities of biodiversity monitoring based on Hungarian light trap networks. *Applied Ecology And Environmental Research* 8(3): 223-239.

8. További tudományos közlemények

- [1] Béres Cs. Z. – Bornemisza I. – Gimesi L. (2002): Hunveyor 2 (Gyakorló űrszonda építése és vezérlése a Pécsi Tudományegyetemen); „Tavaszi szél 2002” nemzetközi, tudományos konferencia (poszter), Gödöllő.
- [2] Béres Cs. Z. – Bornemisza I. – Gimesi L. (2002): Térinformatika a rekultivációban; „Tavaszi szél 2002” nemzetközi, tudományos konferencia (poster), Gödöllő.
- [3] Béres Cs. Z. – Bornemisza I. – Gimesi L. (2002): Felszínmozgási monitoring; „Tavaszi szél 2002” nemzetközi, tudományos konferencia (poszter), Gödöllő.
- [4] Béres Cs. Z. – Bornemisza I. – Gimesi L. (2002): Komplex térinformatikai rendszermodellek kidolgozása, Előadás a „I. Pécsi Hulladékgazdálkodási és Tájrehabilitáció” konferencián, Pécs.
- [5] Béres Cs. Z. – Bornemisza I. – Gimesi L. (2002): Bányászati tevékenység által érintett bányatelkek rekultivációjának monitoring rendszere; Poszter bemutató és szekció előadás „A térinformatika szerepe az agrárstruktúra átalakításában és a vidékfejlesztésben” konferencián, Kaposvár.
- [6] Diósy, T., Roskó, F. Hegyi, S., Keresztesi, M., Kovács, B., Gránicz, K., Drommer, B., Tóth, Sz., Cech, V., Béres, Cs. Z., Gimesi, L., Herbert, J., Imrek, Gy., Lengyel, P., Febriczy, A., Bérczi, B. (2000): New Instrument Assemblages for Planetary Geology on the Hunveyor-1 & 2: Experiment university landers in Hungary: of the Eötvös and Janus Pannonius universities; National Institute of Polar Research, Tokio, Yamaguchi University; (adsabs.harvard.edu/cgi-bin/nph-bib_query?bibcode=2000anme), (yamato.nipr.ac.jp/AMRC/symposium/2000/list_2000.pdf)
- [7] Drégelyi-Kiss, Á., Gimesi, L., Hufnagel, L. (2010): Examination the interaction between the composition of a theoretical ecosystem and the increase in the atmospherical CO₂ level. Hungarian Journal of Industrial Chemistry, 38(2): 201-206
- [8] Ferenczy, A., Varga, R. D., Bíró, I., Kovács, A., Petrányi, G., Hirka, A., Szabóky, Cs., Isépy, I., Priszter, Sz., Türe, D., Gimesi, L., Hufnagel, L. (2010): Fenológiai jelenségek és meteorológiai indikátorok kapcsolatának összehasonlító elemzése rovar és növény adatsorok alapján. LII. Georgikon napok, Nemzetközi tudományos konferencia, Keszthely.
- [9] Ferenczy, A., Eppich, B., Varga, R. D., Bíró, I., Kovács, A., Petrányi, G., Hirka, A., Szabóki, Cs., Isépy, I., Priszter, Sz., Türe, D., Gimesi, L., Garamvölgyi Á., Homoródi, R., Hufnagel, L. (2010): Comparative analysis of the relationship between phenological phenomena and meteorological indicators based on insect and plant monitoring. Applied Ecology and Environmental Research 8(4): 367-376.
- [10] Gimesi László (1982): Rajzgépek az automatizált műszaki tervezésben; Előadás az Országos Automatizált Műszaki Tervezési Konferencián, Budapest.
- [11] Gimesi László (1983): Rajzgépek alkalmazása a tervezésben és gyártásban. Automatizálás, 3: 28-32.
- [12] Gimesi László (1984): Interaktív műszaki tervezés; PÉCSITERV Számítástechnikai tájékoztató, 4.

- [13] Gimesi László (1985): Hatékonyság növelése a műszaki tervezésben; PÉCSITERV Számítástechnikai tájékoztató, 5.
- [14] Gimesi László (1984): Hagyományos számítástechnikai eszközök alkalmazása a műszaki tervezésben; Előadás a „Számítástechnika az építőiparban” konferencián, Budapest.
- [15] Gimesi László (1985): Rajzgépek szoftver kompatibilitásának megvalósítása. Automatizálás, 8: 31-37.
- [16] Gimesi László – Meskó András (1989): Számítástechnika alkalmazása az építéstervezésben, Pécsi Műszaki Szemle, 34(1-4): 11-13.
- [17] Gimesi László (1999): Számítógépes bűnözés szakértői szemmel; Előadás a PÉCSINFÓ kiállításon, Pécs.
- [18] Gimesi László (2000): Multimédia szabványok; (Internetes távoktatási tananyagfejlesztés), PTE Regionális Oktatási központ Pécs.
- [19] Gimesi László (2001): Multimédia hardvereszközök és szabványok; (Internetes távoktatási tananyagfejlesztés), PTE Regionális Oktatási központ Pécs.
- [20] Gimesi László (2002): Digitális terepmodell a tájrendezésben, Geográfus Doktoranduszok VII. Országos Konferenciája, (szekció előadás, kiadvány), ELTE TTK.
- [21] Gimesi László (2002): Információfeldolgozás és a digitális terepmodell (szekció előadás, kiadvány), Számítástechnikai konferencia, Kolozsvár.
- [22] Gimesi László (2002): Térinformatika az űrkutatásban (Gyakorló űrszonda építése a Pécsi Tudományegyetemen) (szekció előadás), XI. Térinformatika az oktatásban konferencia, Szent István Egyetem.
- [23] Gimesi László (2002): Hunveyor-2 űrszonda az oktatásban (poszter), XI. Térinformatika az oktatásban konferencia, Szent István Egyetem.
- [24] Gimesi László (2003): Bevezetés az informatikába (egyetemi jegyzet), Pécsi Tudományegyetem, Krónika, Pécs.
- [25] Gimesi László (2004): IBM PC számítógépek programozása assembly nyelven (egyetemi jegyzet), Pécsi Tudományegyetem.
- [26] Gimesi, L., Béres, Cs. Z., Bérczi, Sz., Hegyi, S., Cech, V. (2004): Teaching planetary GIS by constructing its model for the test terrain of the Hunveyor and Hussar. 35th Lunar and Planetary Science Conference. Houston. Texas.
- [27] Gimesi L. (2004): Mesterséges intelligencia alkalmazása a rekultivációban; Előadás a II. Alkalmazott Informatika konferencián, Kaposvár.
- [28] Gimesi László (2005): Talajvíz ionkoncentrációjának vizsgálata Előadás a IV. Alkalmazott Informatika konferencián, Kaposvár.
- [29] Gimesi László (2006): A csapadékváltozás tendenciái Magyarországon; Előadás az V. Alkalmazott Informatika konferencián, Kaposvár.
- [30] Gimesi L. (2006): Csapadék változása Magyarországon a XX. században; Előadás a II. Kárpát-medencei környezettudományi konferencián; Pécs.
- [31] Gimesi László (2006): Talajvíz ionkoncentrációjának vizsgálata (Közelítő módszerek összehasonlítása); Acta Agraria Kaposváriensis, 10(1): 53-63.

- [32] Hegyi, S, Kovács, B., Keresztesi, M., Béres, Cs. Z., Gimesi, G., Imrek, Gy., Lengyel, P., Herbert, J. (2000): Experiments on the planetary lander station and on its rover units of the Janus Pannonius University, Pécs, Hungary; Lunar and Planetary Science, Huston, (adsabs.harvard.edu/cgi-bin/nph-bib_query?bibcode=2000LPI).
- [33] Hegyi, S., Kovács, B., Keresztesi, M., Gimesi, L., Halász, A., Bérczi, Sz. (2003): New gas sensing device to the Hunveyor-2 university lander of the Pécs University, Hungary, 34th Lunar and Planetary Science Conference, Houston, Texas.
- [34] Hufnagel – Sipkay – Drégely-Kis – Farkas – Türei – Gergócs – Petrányi – Baksa – Gimesi – Eppich – Dede – Horváth (2008): Klímaváltozás, biodiverzitás és közösségökológiai folyamatok kölcsönhatásai, in: Harnos – Csete: Klímaváltozás: Környezet – Kockázat – Társadalom, Szaktudás Kiadó Ház, Budapest: 229-267
- [35] Kúti, Zs., Hirka, A., Petrányi, G., Szabóky, Cs., Gimesi, L., Hufnagel, L., Ladányi M. (2010): A kis téliaraszoló (*Operophtera brumata* L.) aktivitásának modellezése abiotikus paraméterekkel. *Agrárinformatika / Agricultural Informatics*, 1(1): 40-46.
- [36] Szeidl L. – Gimesi L. – Markó T. (2005): Időjárás generátor és a sztochasztikus modellezés néhány kérdése; Plenáris előadás a VII. Magyar biometriai és biomatematikai konferencián; Budapesti Corvinus Egyetem, Budai Campus.
- [37] Varga, B., Boldizsár, A., Gimesi, L. (2007): Some Remarks on Water Balance Components of Lake Balaton; *Georgikon for Agriculture*, 10(1): 69-82.
- [38] Verasztó, Cs., Kiss, K. T., Sipkay, Cs., Gimesi, L., Vadadi-fülöp, Cs., Türei, D., Hufnagel, L. (2010): Long-term dynamic patterns and diversity of phytoplankton communities in a large eutrophic river (the case of river Danube, Hungary), *Applied Ecology and Environmental Research*, 8(4): 329-349.