istvan60_104_23

A hidrológiai ciklus különböző elemeiből származó környezeti adatok vizsgálata idősoros és többváltozós geomatematikai módszerekkel

MTA doktori rövid értekezés

Készítette:

Hatvani István Gábor
HUN-REN Csillagászati és Földtudományi Kutatóközpont, Földtani és Geokémiai Intézet

Budapest, 2024

Tartalom

Rövidítések jegyzéke[1]

AIC	Akaike információs kritérium
AT, HU	Ausztria, Magyarország ISO 3166-1 alpha-2 kódja
BIC	Bayesian információs kritérium
BOI_5	Ötnapos biológiai oxigénigény
CCDA	Kombinált klaszter- és diszkriminanciaanalízis
Chl-a	Klorofill-a
D	Tengertől való távolság (km)
δ_p	Csapadék stabilizotóp-összetétel (‰)
DFA	Dinamikus faktoranalízis
ELE	Tengerszint feletti magasság (m)
EU	Európai Unió
GNIP	Global Network for Isotopes in Precipitation
HCA	Hierarchikus klaszteranalízis
IDW	Távolsággal fordítottan arányos súlyozás (inverse distance weighting)
KBVR	Kis-Balaton Vízvédelmi Rendszer
KOI	Kémiai oxigénigény
LAT	Földrajzi szélesség
LCsVV	Lokális csapadékvízvonal
LDA	Lineáris diszkriminanciaanalízis
LON	Földrajzi hosszúság
MAE	Átlagos abszolút hiba
mvp	mintavételi pont
NYUDUVIZIG	Nyugat-dunántúli Vízügyi Igazgatóság
Old P	Oldott foszfor
Oldott O	Oldott oxigén
ORP	Oldott reaktív foszfor
Ö Leba	Összes lebegőanyag
ÖN	Összes nitrogén
ÖP	Összes foszfor
PCA	Főkomponens-analízis
PP	Partikulált foszfor
Q	Vízhozam
$R^{\sim 2}$	Korrigált determinációs együttható
RF	Véletlen erdő (random forest)
RMA	Reduced major axis
RMSE	Átlagos négyzetes hiba gyöke
SFIB	Sztenderd fekáliás indikátorbaktérium
SVM	Support vector machine
Szerves N	Szerves nitrogén
VIF	Variancia inflációs faktor
WCA	Wavelet-koherenciaanalízis
WMMWL	Western Mediterranean Meteroric Water Line (Nyugat-mediterrán Csapadékvízvonal) (Celle-Jeanton et al., 2001)
WW-P	Szennyvíz-foszfortartalom
δ^2H és $\delta^{18}O$	Hidrogén- és oxigénizotóp-összetétel (‰)

[1] Az értekezésben, ahol releváns, Demény, A., 2003. Stabilizotóp-geokémia és termometria: Hogyan és mire? Földtani Közlöny 133(2): 263-270 által javasolt magyar geokémiai nevezéktant és szóhasználatot követtem.

Dolgozatom néhai Szüleimnek és Nagymamámnak dedikálom, köszönök mindent, amit tőletek kaptam...

1. BEVEZETÉS

A víz átfogóan meghatározza az élő és élettelen környezet folyamatait. Alapvető környezeti erőforrás, mely fenntartja az életet, az ökoszisztémákat és az emberi társadalmat is. A globális vízkörzés a víz folyamatos mozgását írja le a földfelszínen, a földfelszín felett és alatt. A víz széles skálájú oldószerként szállító funkciót is ellát: folyamatos, a napenergia által hajtott mozgása biztosítja az élet fenntartásához létfontosságú kémiai elemek vándorlását a hidroszféra, az atmoszféra, a bioszféra és a litoszféra között (Robertson et al., 2022).

A modern társadalom nagymértékben függ az éghajlat által meghatározott szárazföldi vízkörforgástól (Vörösmarty és Sahagian, 2000), így elkerülhetetlen az antropogén tevékenység által a hidrológiai ciklusban okozott változások visszacsatolása magára az ember környezetátalakító tevékenységére és egyúttal természeti környezetünkre (Vörösmarty et al., 2010). A XX. század második felétől egyre markánsabban körvonalazódott az emberi tevékenység hatása a globális vízkörforgásra (Abbott et al., 2019; Bindoff et al., 2013). Nagy általánosságban elmondható, hogy globális skálán pl. a légszennyezés bizonyítottan lassította a vízkörforgást a XX. sz. közepén és annak második felében. Az ezredfordulóra azonban az antropogén aeroszol-kibocsátás csökkentésére irányuló sikeres törekvések hatására már mindinkább az üvegházhatású gázok kibocsátása dominálta az antropogén sugárzási kényszert, mely a vízkörzés felgyorsulását eredményezte; ami a közeljövőben a csapadékesemények intenzitásának további növekedését vonhatja maga után (Wu et al., 2013).

Jelen munka az (édes)vízkészletekre fókuszál, melyek fokozottan ki vannak téve az ember környezetmódosító hatásának (Vörösmarty és Sahagian, 2000). Az édesvíz emberi felhasználásra is fenntarthatóan hozzáférhető része – elsősorban felszíni és felszín alatti nem fagyott víz – nagyjából 5 000 és 9 000 km^3 $év^{-1}$ között mozog, ami megközelíti a tényleges becsült éves globális vízfelhasználás mértékét (3 800-tól 5 5000 km^3 $év^{-1}$) (Abbott et al., 2019). A második legjelentősebb édesvízkészlet a jégsapkákban és a permafrosztban eltárolódott víz, amely a globális vízkészlet közel 2%-a (Abbott et al., 2019; Trenberth et al., 2007).

A fent említett globális skálán ható antropogén eredetű tényezők mellett ki kell emelni néhány – (makro)regionális skálán jelentkező – antropogén tényezőt, melyek mind közvetlenül befolyásolják a hozzáférhető édesvízkészlet egyre csökkenő mennyiségét (Mekonnen és Hoekstra, 2016) és eloszlását (Döll et al., 2012). Ilyenek például a folyószabályozások (Döll et al., 2009), az állóvizek lecsapolása, az ipari és mezőgazdasági célú felszín alatti vízkivétel (Döll és Siebert, 2002). Az urbanizáció, az iparosodás és a népességnövekedés elkerülhetetlen velejárója a vízforrások minőségének degradálódása pl. diffúz mezőgazdasági (Donohue et al., 2005; Hatvani et al., 2022) vagy pontforrás eredetű vízszennyezések által (Csathó et al., 2007; Hatvani et al., 2018b; Sisák, 1993).

Számtalan nemzetközi, akár kontinenseken is átívelő szabályozás létezik (Faragó, 2014), amelyek irányt mutatnak, hogy vizeinket fenntartható módon használjuk, így megfelelő minőségben, mennyiségben adhassuk át a jövő generációi számára. Az 1990-es évek végén kezdte el kialakítani az Európai Unió (EU) elődje a Vízkeretirányelvet, amely ajánlás, pontosabban szabályozási rendszer azzal a céllal jött létre, hogy 2015-re minden EU-s tagországban „jó (ökológiai) állapotba" kerüljenek felszíni vizeink és vízi környezeteink (EC, 2000). Ennek felszín alatti vizekre vonatkozó megfelelőjét 2006-ban fogadta el az Európai Bizottság (EC, 2006b). Az Amerikai Egyesült Államokban a Vízkeretirányelvvel „egyenrangú" szabályozást már közel fél évszázaddal korábban bevezették, ez az ún. 'Clean Water Act' (US Government, 1948). Az Európai Unió kifejezetten rekreációs célzattal is igénybe vett vizekre vonatkozó minőségi elvárásait és javaslatait egy erre dedikált irányelve (’Bathing Water Directive’) szabályozza (EC, 2006a).

Amellett, hogy az édesvízkészleteink mennyiségét és hozzáférhetőségét negatívan befolyásoló folyamatok erősödése aggodalomra ad okot. Azonban az elmúlt évtizedek problémaközpontú technológiai fejlesztésinek hála, egyre növekvő mennyiségben férünk hozzá a hidrológiai ciklus főbb tározóit és a köztük zajló anyagáramlást (Abbott et al. (2019): 3. ábra)[2] leíró adatokhoz, amelyek segíthetnek jobban megérteni a globális léptékű hidrológiai folyamatokat és azoknak a hidrológiai ciklus által érintett szférákra gyakorolt (kölcsön)hatását (Koutsoyiannis, 2020). Feladatunk ezek megfelelő szintetizálása, feldolgozása, kiértékelése modern geomatematikai eszközökkel úgy, hogy az a lehető legjobban írja le a vizsgált jelenségeket; és feladat egyúttal az így kapott tudás megosztása. A geomatematika egy interdiszciplináris tudományág, melynek célja a természeti, társadalmi és gazdasági tényezők által befolyásolt komplex és dinamikus földi rendszer megértésének elősegítése matematikai módszerek alkalmazásával: a leíró gondolkodás, megfelelő egyszerűsítések és matematikai koncepciók kidolgozásának eszközeivel (Freeden, 2010).

Dolgozatomban pontosan a geomatematika céljai mentén és eszköztárának alkalmazásával (Freeden (2010): 1. ábra) hozok ilyen gyakorlatokra példákat, kiemelve a felszíni vizek, és ezen belül a (i) sekély tavak trofikus állapotának és (ii) esetleges fekáliás szennyezésének kérdéskörét, melyre a következő (1.2.) alfejezetben térek ki részletesebben. Továbbá (iii–iv) betekintést adok azon izotóphidrológiai eszközökbe (csapadék stabilizotóp-összetétel; Gat et al. (2001)), melyekkel hatékonyan nyomon követhetjük a vízmolekula útját a vízkörforgásban a csapadékhullástól a felszíni és felszín alatti vizekig, majd az ivóvízkészletekig (Vreča és Kern, 2020) (1.3. fejezet).

1.1. Környezeti adatok tér- és időbeli változékonysága, hozzáférhetősége – gyakorlati szempontok

A hidrológiai ciklus elemeiben bekövetkezett természetes és antropogén eredetű változások észlelése, nyomon követése, esetleges enyhítése és előrejelzése az egyik legégetőbb problémája a mai környezet- és vízvédelemnek, legyen szó általános vízminőségromlásról, vagy valamely kifejezett paraméterkörhöz köthető változásról.

A mintavételezési (pl. távérzékelés, kihelyezhető mérőeszközök) és analitikai technológiák fejlődése lehetővé tette ezen eszközök egyre szélesebb kőrű elérhetőségét és alkalmazhatóságát. Így exponenciális növekedésnek indult a vízi környezetekből származó adatok mennyisége pl. Read et al. (2017). Ha a távérzékeléssel nyert adatokból készült, a kontinentális vizek minőségével foglalkozó tanulmányok számát nézzük, a 2000-es évek elején évente átlagosan 20-25 tanulmány készült, míg ez a szám a 2010–2020-as időszakra megnégyszereződött (Topp et al., 2020). Meg kell azonban jegyezni, hogy az adatok mennyiségi növekedése önmagában nem elégséges, azoknak reprezentatívaknak kell lenniük az adott közegre, és a paraméterkörnek relevánsnak kell lenni a feltett kérdések megválaszolásához (**1.2-1. ábra**).

E kritériumok közül a dolgozatban bemutatott tanulmányok szempontjából a legfontosabbak:

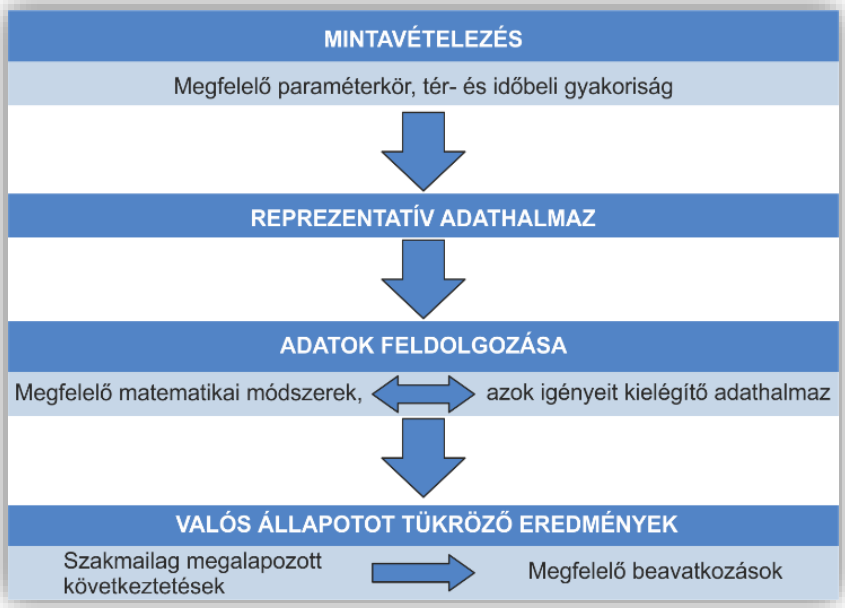
- Egy adott változóra térben legyen annyi adatpont, hogy azok között legyen megfelelő erősségű a térbeli autokorreláció, i.e. kapcsolatiságot lehessen vizsgálni közöttük

[2] A hivatkozások utáni kettőspontot követő ábraszámok az adott tanulmány ábrájára hivatkoznak, és nem a jelen dolgozatra. A dolgozat alapjául szolgáló tanulmányok különlenyomatai a dolgozat végén találhatók.

(Webster és Oliver, 2008); ha nem egy egyszeri mintavételi kampány adataival kell foglalkoznunk, akkor

- minél hosszabb időszakot fedjenek le a mért változók idősorai (Burt et al., 2014), és rendelkezzenek megfelelő időbeli adatsűrűséggel;
- időben lehetőleg folytonos legyen az adatlefedettség, hiszen a jelentős „szakadások” pótlása nem, vagy csak jelentős bizonytalansággal lehetséges (Van Buuren, 2018).
- Ha pusztán szakaszos adathiányok jönnek létre, azok ne okozzák az idősorok „szakadását”, hanem „csak” az egyenközűséget változtassák meg (pl. kétheti, havi, heti mintavételezési gyakoriság váltakozik). Ehhez:
  1. *Álljanak rendelkezésre további változók*, melyekkel azok pótolhatók/becsülhetők lehessenek regresszióval vagy egyéb módszerekkel: periodicitásvizsgálat és regresszió kombinációjával, pl. Kovács et al. (2017), neurális hálózatokkal, pl. Csábrági et al. (2019), vagy egyéb eszközökkel (Van Buuren, 2018).
  2. *Ha azonban nem áll rendelkezésre egyéb segédváltozó*, és egyenközűvé kell tenni az adatokat, akkor (i) legyen értelme aggregálni az idősorokat (átlag, medián vagy egyéb statisztika képzése pl. Hatvani et al. (2011b, 2014a); Kovács et al. (2012b,c, 2015a); Magyar et al. (2013); Tanos et al. (2015)), vagy (ii) az idősorok saját tulajdonságait felhasználva lehessen pusztán matematikai úton azokat pótolni.

Mindazonáltal fontos leszögezni, hogy bármilyen transzformáció, amely „hozzáad” adatot az eredeti idősorhoz vagy „elvesz” belőle, elkerülhetetlenül megváltoztatja annak spektrális tulajdonságait (Schulz és Stattegger, 1997), amit minden esetben meg kell vizsgálni, ha spektrumanalízis a cél (Hatvani et al., 2018a).



1.2-1. ábra: A mintavételezés és a kiértékelés fő lépéseinek sematikus ábrája (Hatvani et al., 2011a).

E fenti szempontok mind meghatározzák azon geomatematikai módszerek körét, amelyek alkalmazhatóak egy adott környezeti probléma kezelésében. E kérdésről egy tanulmányban részletesen értekeztünk kollegákkal (Hatvani et al., 2014b), de ennek taglalása jelen munkának nem célja.

1.2. Eutrofizáció és fekáliás szennyezések

Az elsődleges termelő szervezetek elszaporodása a foszfor- (P) (Schindler, 1974; Schindler et al., 2016; Vitousek et al., 2010) és/vagy nitrogén- (N) tartalmú tápanyagok (Carpenter, 2008) feldúsulásának következtében a vizek trofikus állapotának romlásához vezet, ami kritikus szinten anoxikus állapotokkal (Lau és Lane, 2002), a magasabb rendű szervezetek tömeges pusztulásával és az anaerob baktériumok elszaporodása révén a víz toxicitásának növekedésével jár (Wetzel, 2001). Az elmúlt évtizedekben ez az egyik legkritikusabb folyamattá vált a tavaink vízminőségromlása esetén (Scheffer, 2013; Schindler, 1974; Schindler et al., 2016; Wetzel, 2001). A sekély tavak, mivel (i) jól átkevertek (Chapman (1996)), (ii) viszonylag nagy a vízfelszín/mélység-hányadosuk és (iii) intenzív tó–szárazföld-, levegő–víz- és víz–üledék-interakciók jellemzik (Wetzel, 2001), még inkább kitettek az eutrofizálódásnak, mint a többi felszíni víztest. A kérdés fontosságát jelzi, hogy minden kontinensről számos tanulmány vizsgálta már a sekély tavakban tapasztalt tápanyag-feldúsulással párhuzamosan fokozódó eutrofizálódás kérdéskörét, pl. Ázsia (Fink et al., 2018; Qin et al., 2007), Európa (Fink et al., 2018; Hatvani et al., 2014a, 2020a; Herodek et al., 1982; Istvánovics et al., 2007; Jeppesen et al., 2007; Padisák és Reynolds, 2003; Sebestyén et al., 2019; Somlyódy et al., 1983; Somlyódy és van Straten, 1986; Van der Molen és Portielje, 1999), Észak-Amerika (Fink et al., 2018; Oberholster et al., 2006), Közép-Amerika (López-López et al., 2016), Dél-Amerika (Fink et al., 2018), Afrika (Fink et al., 2018; Muli, 1996) és az Antarktisz tavai esetében (Izaguirre et al., 2021). A korábbi példák közül kiemelem a Balatont, Közép-Európa legnagyobb sekély vizű tavát (Padisák, 1985). A Balaton a XX. sz. második felében jelentős eutrofizálódáson, ennek kiküszöbölésére különféle beavatkozásokon, majd oligotrofizáción ment keresztül (Hatvani et al., 2020a; Herodek, 1984; Istvánovics et al., 2007); ennek részletes vizsgálata jelen dolgozat egyik fő eleme.

A dolgozat szempontjából egy másik fontos tényező a fekáliás szennyezések kérdésköre. A XX. sz. második felére a rekreációs célra egyre inkább használt mérsékelt övi tavak (Cooper, 2006) vízminősége jelentősen leromlott, ahogy ezt fentebb is említettem, és ezzel egyidejűleg ökoszisztéma-szolgáltatásaik a rekreációs célra történő használatuk lehetőségében szűkültek (Dokulil, 2014). Számos intézkedés, mint pl. a szennyvízelvezetés, csatornázás (Hatvani et al., 2014a; Somlyódy, 2018), a foszfortartalmú mosószerek fokozatos kivezetése (EC, 2007; Litke, 1999), vagy egyes, tájvédelmi körzetek létrehozásával kapcsolatos lépések (Parry, 1998), jelentős javulást eredményeztek a tavak vízminőségében. Noha a fekáliás szennyeződés az egyik elsődleges környezetszennyező a világon (Santo Domingo et al., 2007), melyre már a XIX. századból is rendelkezésre állnak jól dokumentált esetek (Somlyódy, 2018), ennek vonatkozásában csak 2006-ra született EU-s iránymutatás, az EU 'Bathing Water Directive' (EC, 2006a) melynek célja, hogy rendszeres vízminőség-ellenőrzés (monitoring) és üzemvezetési–technológiai gyakorlatok segítségével minimalizálja a lakosság fürdővizek rekreációs használatából eredő egészségügyi kockázatát (Bedri et al., 2016). Sajnálatosan a kommunális szennyvíz továbbra is gyakori okozója maradt a fürdőhelyek fekáliás szennyeződéseinek (de Brauwere et al., 2014) a világ minden táján: pl. Észak-Amerikában (Nevers et al., 2014), Európában (Hatvani et al., 2018b; Poté et al., 2009), Afrikában (Leboulanger et al., 2021), Ázsiában (Abdulaziz et al., 2023; Kongprajug et al., 2021; Shiekh et al., 2006), Ausztráliában (Leeming et al., 1997; Schang et al., 2016). Akárcsak az eutrofizáció kezelése, ez is kulcsfontosságú felszíni vizeink jó minőségének megőrzése céljából, és e probléma vizsgálatához és enyhítéséhez/megoldásához minél hosszabb és az eutrofizációt átfogóan jellemző adatsorok szükségesek.

1.3. A hidrológiai ciklus mintavételezésének kezdetei és a paleoproxi-adatok szerepe: fókuszban a csapadék stabilizotóp-összetétele

A hosszú időtartamot lefedő és nagy időbeli sűrűségű műszeres mérések elterjedésével növekvő adatmennyiség nem alkalmas az akár évszázadokkal ezelőtti környezeti állapotok/folyamatok leírására, nemhogy globális, de regionális skálán sem. A leghosszabb, közvetlenül is a hidrológiai ciklushoz köthető, máig folyamatosan üzemelő mintavételi hálózat az Osztrák Meteorológiai és Geodinamikai Intézet (ZAMG), amely 1851-es alapítása óta szolgáltatja a meteorológiai adatokat. Hazánkban 1870 óta végez rendszeres adatgyűjtést az eredeti nevén Meteorológiai és Földdelejességi Magyar Királyi Központi Intézet, a mai Országos Meteorológiai Szolgálat elődje (Antal, 1996-2000). Ahogyan időben haladunk a múlt felé, egyre kevesebb ponton és egyre kevesebb változóról van közvetlen műszeres adatunk. Ahhoz, hogy (i) a hidrológiai ciklusban és annak elemeiben el tudjuk különíteni az antropogén hatások okozta változásokat a természetes változékonyságtól, illetve, hogy (ii) előrejelzéseket tudunk végezni a jövőre vonatkozóan, (iii) meg kell, hogy értsük a múltban – akár geológiai időskálán – bekövetkezett hidrológiai változásokat, legyen szó pl. vízgyűjtő- (Korponai et al., 2010; Saito et al., 2008), vagy kontinens méretű folyamatokról, lásd pl. a csapadék stabilizotóp-összetétel (δ_P) tér és időbeli változékonyságának modellezését az Antarktiszon (Goursaud et al., 2018).

Amennyiben a közvetlen műszeres méréseket megelőző idők folyamatait kívánjuk megérteni, akkor közvetett, ún. „proxi”-adatsorok segítségével szerezhetünk információt elmúlt korok (hidro)klímájáról, azok változékonyságáról, pl. cseppkövek, jégfuratok, mélytengeri üledékekben rögzült környezeti változók (pl. csapadék stabilizotóp-összetétel) rekordjaival (NRC, 1996). A paleoklimatológiában a proxiadatsorok elmúlt korokban uralkodó környezeti állapotok természetes archívumai, melyek akár geológiai időskálán is képesek tükrözni a rétegeik keletkezésekor uralkodó környezeti állapotok lenyomatát a bennük rögzített paraméterek (pl. izotóparányok) alapján (Ouellet-Bernier és de Vernal, 2018). Jelen dolgozat szempontjából az üledékes (többnyire folyamatos és jellemzően nemlineáris képződésű) paleoproxik közül is a jégfuratok, és az azokban rögzült δ_P kap kiemelt szerepet a múltbéli hidrológiai viszonyok rekonstrukciójában betöltött kiemelt szerepük miatt (Rozanski et al., 1997; Steiger et al., 2017).

A csapadék stabilizotóp-összetétel egy adott elem nehéz és könnyebb izotópjának arányát fejezi ki egy nemzetközi sztenderdhez viszonyítva ezrelékben (‰), deltával jelölve (Dansgaard, 1964), hidrogénizotóp-összetétel esetében δ^2H, oxigén esetében δ^{18}O. A vízminták esetében a nemzetközi sztenderd, amihez a kapott arányokat viszonyítjuk, az ún. Vienna-Standard Mean Ocean Water (Coplen, 1994).

A δ_P rendkívül konzervatív nyomjelzője a vízkörzésben végbemenő környezetfizikai változásoknak (Dansgaard, 1964). Alkalmazásukkal pontosabban meghatározható különböző vizek eredete, hidrológiai folyamatok mibenléte, mértéke (Bowen et al., 2019; Clark és Fritz, 1997; Gat et al., 2001). A párolgás, a kondenzáció és egyéb fázisváltások során a nehezebb izotóp frakcionálódik a könnyebb izotóptól, ezzel rögzítve a páratömegek környezeti lenyomatát. A δ_P ilyen formán kulcsfontosságú információt adhat a páraforrás eredetéről (pl. Gat et al. (2003)), a páratömegek által megtett útvonalakról (Gat et al., 2001), a csapadékhullás körülményeiről (Araguás-Araguás et al., 2000; Dansgaard, 1964; Merlivat és Jouzel, 1979) egészen a felszínben végbement folyamatokig.

A csapadékból mért δ^2H és δ^{18}O kapcsolatát leíró összefüggést csapadékvízvonalnak nevezzük (lásd pl. Putman et al. (2019): 5. ábra), és egy alapvető eszköze a hidrológiai folyamatok vizsgálatának (Fórizs, 2005; Gat, 2005). Általános gyakorlat, hogy egy adott helyre lokális csapadékvízvonalat (LCsVV) határoznak meg. Az összes édesvízi δ_P legjobb lineáris illesztése a globális csapadékvízvonalat adja meg (Craig, 1961), amit később mennyiséggel súlyozott éves átlagok alapján felülvizsgáltak így megkapva a ma is használt empirikus összefüggést: $\delta^2H = 8,20\ (\pm\ 0,07) \times \delta^{18}O + 11,27\ (\pm\ 0,65)$ (Rozanski et al., 1993). A globális csapadékvízvonal gyakorlatilag az LCsVV-k sokaságának tekinthető, amelyek regionális vagy helyi léptékben meredekségük és tengelymetszetük szempontjából nagymértékben eltérnek (Kendall és Coplen, 2001; Sharp, 2017).

2. A DISSZERTÁCIÓ MOTIVÁCIÓI

Megállapítható, hogy:

(i) bár a tudományos társadalom számára egyre nagyobb mennyiségű nyers adat áll rendelkezésére, az elemzésükből, szintézisükből kinyerhető tudástöbblet növekedése elmarad attól a szinttől, amely az adatbőség okán elvárható lenne,

(ii) nem elegendő „pusztán statisztikusi" szemmel megvizsgálni az adatokat, hanem tisztában kell lenni azok eredetével és környezeti jellegükből adódó kapcsolati viszonyaikkal, hogy a lehető leghatékonyabban nyerhessük ki a szükséges információt,

(iii) az egyre növekvő számú mintavételi pontból létrejövő hidrológiai mintavételi hálózatokat rendszeresen felül kell vizsgálni a redundancia és szükséges lefedettség hiánya szempontjából (pl. Hatvani et al. (2017b, 2018b, 2020b, 2021); Kardos és Clement (2020); Kovács et al. (2014a, 2015a, 2012a); Tanos et al. (2015)) szem előtt tartva a pontok egyedi szerepét is a hálózatban (Innes, 1998). Ez elméleti és gyakorlati szempontból is egyaránt lényeges feladat, mivel redundancia feltárása estén többletkiadásokat lehet lefaragni bármilyen környezeti mintavételi hálózat esetén pl. Nunes et al. (2004).

Mindezek fényében a dolgozat átfogó motivációja, hogy egységes rendszerbe foglalt geomatematikai eszköztár (idősoros, többváltozós és geostatisztikai módszerek) alkalmazásával példákat hozzon a hidrológiai ciklus terresztriális felszíni víztestjeit érintő (i) általános vízminőségi és trofikus állapotváltozásainak (Hatvani et al., 2015, 2017a, 2020a), és (ii) fekáliás szennyeződési forrópontjainak feltárására (Hatvani et al., 2018b; Herzig et al., 2019). Továbbá, hogy bemutassa a (iii) modern csapadék stabilizotóp-összetételéből számított csapadékvízvonal meredekségének és tengelymetszetének térbeli becslését a tágabb értelemben vett mediterrán térségre (Hatvani et al., 2023) és végül (iv) egy izotópos tájkép létrehozását klasszikus geostatisztikai módszerekkel, amely hozzájárul a múltbéli páraáramlási útvonalak feltárásához és a nyugat-antarktiszi jégfurathálózat optimalizálásához (Hatvani et al., 2017b). Ezek az eredmények számos ponton kapcsolódnak a hidrológia kurrens kutatási kérdéseihez (Blöschl et al., 2019). A különböző kutatásokhoz kapcsolódó specifikus célokat az adott fejezetekben taglalom a vizsgált területek és hozzájuk kapcsolódó tudományos kérdések bemutatása után.

3. ADATOK ÉS MÓDSZEREK

Mivel az értekezés a hidrológiai ciklus különböző szféraiból származó adatok geomatematikai feldolgozására hoz példákat, ezért a tárgyalt környezeti adatok is jelentős diverzitást mutatnak. A vizsgált területek körét három felszíni víztest, a Zala (**4.1-1A ábra)** és hozzá kapcsolódva Kis-Balaton Vízvédelmi Rendszer (**4.1-1B ábra**), a Balaton (**4.1-1C ábra**) (4.1.1. fejezet) és a Fertő (**4.2-1C ábra;** 4.2.1. fejezet), továbbá a tágabb értelemben vett mediterrán térség (továbbiakban Mediterráneum) δ_P mintavételi hálózata (**4.3-1. ábra;** 4.3.1. fejezet), végül egy nyugat-antarktiszi jégfurathálózat idősorai alkotják (**4.4-1. ábra;** 4.4.1. fejezet). Először a mintaterületekről származó adatok körét (3.1. fejezet) és az ezekre alkalmazott, egységes rendszerbe foglalt adatelemző apparátust írom le általánosan (3.2. fejezet), amelyből egy-egy módszer lényegi jellemzőjét alaposabban is kifejtem (3.2.1. fejezet).

3.1. A vizsgált adatok

Az értekezésben bemutatott tanulmányokhoz kapcsolódó adatok száma közel 300 000, nem számolva a függetlennek tekintett magyarázó és egyéb prediktor változók értékeivel (**3.1-1. táblázat**). Ezen adatmennyiség kb. 75%-a adta a dolgozatban bemutatott eredmények alapját. A pontos részleteket az eredeti tanulmányok tartalmazzák: Hatvani et al. (2015, 2017a,b, 2018b, 2020a, 2023). A felszíni víztestek esetében figyelembe vett vízminőségi változók jellemzőit a dolgozat nem tárgyalja, viszont a csapadék stabilizotópos összetételét leíró paramétereket és a csapadékvízvonalak koncepcióját az 1.3-as fejezetben bevezettem és részletesebben a 3.2.1-es fejezet végén taglalom.

Általánosságban elmondható, azon túl, hogy a bemutatott tanulmányok mind a hidrológiai ciklus egy-egy eleméből származó adatokat dolgoznak fel, a vizsgált változók körében és/vagy a feldolgozásukhoz alkalmazott módszertanban is szorosan összekapcsolódnak. Ez már a **3.1-1. táblázat**ból is kitűnik, és a 3.2-es fejezetben egy egységes megközelítés segítségével, az esettanulmányokat a bennük alkalmazott módszereken keresztül is összekötöm.

A Balaton esetében az adatok éves átlagai képezték a vizsgálatok tárgyát, a KBVR-nél éves átlagok voltak a dinamikus faktoranalízis (DFA) bemenő adatai, és napi adatok a Wavelet-koherenciaanalízis (WCA) esetében, míg a Fertőn minden év március 1–október 31-ig tartó időszak átlagait vizsgáltam évente. A külső magyarázó változókat is ennek megfelelően aggregáltam (**3.1-1. táblázat**).

A Mediterráneum modern (2000–2015), állomásonkénti LCsVV-inek meghatározásánál (lásd 4.3. fejezet) havi felbontású δ_P-idősorok voltak a kezdeti bementi adatok, melyek elsősorban a GNIP adatbázisából (IAEA, 2019) és további irodalmi forrásokból származtak; Hatvani et al. (2023): 2.2. fejezet[3]. Ezek szigorú és alapos adatelőkészítésen mentek keresztül, lásd Hatvani et al. (2023): 2.3.1. fejezet. Ennek első lépése egy vizsgált időtartamra vonatkozó adatmennyiséghez kötött állomáskategorizálás volt (Hatvani et al., 2023; Putman et al., 2019), amelyet Moran I statisztika (Anselin, 1996; Moran, 1948) alapú adatszűrés követett. Amennyiben kiugró értéket véltem felfedezni, az adott állomás d-többletét (Dansgaard, 1964) is megvizsgáltam. Ha az is kiugró értéket mutatott az állomás szomszédaihoz képest (pl. differencia > 10), az adott δ_P-értékeket elhagytam. Ez a megközelítés összetettebb, mint a korábbi tanulmányokban alkalmazott eljárás, amely egy statikus d-többlet-határértéket használt (Bowen, 2008; Nelson et al., 2021) a térbeli összefüggések figyelembevétele nélkül. Az eredmény, hogy egy adott időhorizontban a környezetükhöz képest – kapcsolati viszonyukat tekintve – kiugró értékektől mentes lett a vizsgált adathalmaz. Végeredményképpen 42 db 1. adatminőségi kategóriába tartozó, 20 db 2. kategóriába tartozó állomást használtam fel, melyek közül 11-et összevontam. A kategóriákba kerülés feltételrendszere Hatvani et al. (2023): a 2.3.1-es fejezetében található.

Az antarktiszi jégfuratok esetében fontos kiemelni, hogy az elemzett 60 db jégfurat dokumentált korbizonytalansága ±1 év a Dronning Maud Land (Oerter et al., 2000) és a Ronne Ice Shelf (Graf et al., 1999) területén a vizsgált időszakhoz (1970–1988) legközelebb eső időszakokban. Minden jégfurat esetében az időszakra vonatkozó stabilizotóp-összetétel átlagával (kb. 20 év) számoltam a geostatisztikai vizsgálatok során, így a fent említett korbizonytalanság elhanyagolható volt.

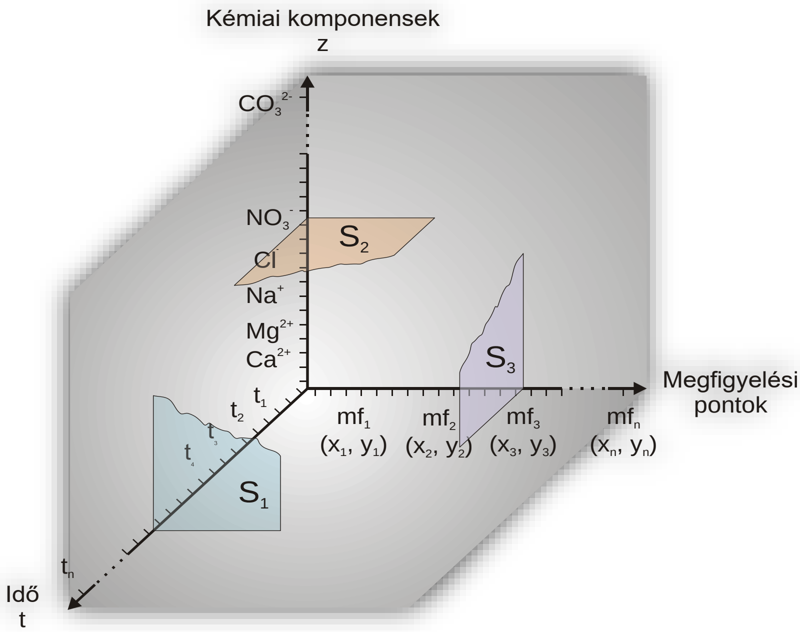
[3] A hivatkozások utáni kettőspontot követő fejezetszámok az adott tanulmány fejezetére hivatkoznak, és nem a jelen dolgozatra.

3.1-1. táblázat. A vizsgált adathalmazok jellemzői: bevont paraméterek, azok által lefedett időszakok, mintavételi pontok, adatok száma, felhasznált módszerek és szoftverek. Az aláhúzott paraméterek (az első oszlop) külső magyarázó változók az adott vizsgálatban, az ezekhez tartozó adatok darabszáma nem tartozik bele a 2. sor értékeibe. Az utolsó előtti két sor prediktor változókat tartalmaz. A táblázatban alkalmazott rövidítések magyarázata a dolgozat 3. oldalán lévő Rövidítések jegyzékében olvasható

<table>
<tr><th>Paraméterek / módszerek</th><th>Mértékegység</th><th>PCA, Geostatisztika</th><th>CCDA, PCA, 'Sen's slope'</th><th>DFA</th><th>WCA</th><th>RMA, IDW, RF, SVM</th><th>Geostatisztika (krigelés)</th></tr>
<tr><td>Víztest/terület (mintavételi pontok száma)</td><td></td><td>Fertő (26 vmvp) (6 meteorológiai pont)</td><td>Balaton (4 vmvp)</td><td colspan="2">Kis-Balaton Vízvédelmi Rendszer (1 ill. 4 vmvp) és griddelt meteo adatok</td><td>Mediterrán régió (249 csapadék mvp)</td><td>Nyugat-Antarktisz (60 jégfurat)</td></tr>
<tr><td>Adat (~db) szűrés előtt</td><td></td><td>108 000</td><td>36 000</td><td>31 000</td><td>117 000</td><td>15 000</td><td>3 000</td></tr>
<tr><td>KOI</td><td rowspan="24">mg l^{-1}</td><td>1992-2013</td><td rowspan="2">1985-2017</td><td rowspan="10">1978-2006</td><td rowspan="10"></td><td rowspan="29"></td><td rowspan="29"></td></tr>
<tr><td>BOI_5</td><td rowspan="11"></td></tr>
<tr><td>Cl^-</td><td rowspan="10"></td></tr>
<tr><td>SO_4^{2-}</td></tr>
<tr><td>HCO_3^-</td></tr>
<tr><td>CO_3^{2-}</td></tr>
<tr><td>Mg^{2+}</td></tr>
<tr><td>Ca^{2+}</td></tr>
<tr><td>Na^+</td></tr>
<tr><td>K^+</td></tr>
<tr><td>ÖN</td><td></td><td>1994-2010</td></tr>
<tr><td>NO_2-N</td><td rowspan="3">1978-2006</td><td rowspan="2"></td></tr>
<tr><td>NH_4-N</td><td rowspan="3">1992-2013</td><td rowspan="4">1985-2017</td></tr>
<tr><td>NO_3-N</td><td rowspan="4">1994-2010</td></tr>
<tr><td>ÖP</td><td></td></tr>
<tr><td>ORP</td><td rowspan="7"></td><td rowspan="5">1978-2006</td></tr>
<tr><td>Ö Leba</td><td rowspan="6"></td></tr>
<tr><td>Fe^{2+}</td><td rowspan="9"></td></tr>
<tr><td>Mn^{2+}</td></tr>
<tr><td>Szerves N</td></tr>
<tr><td>Old P</td><td></td></tr>
<tr><td>PP</td><td rowspan="2">1978-2006</td></tr>
<tr><td>Chl-a</td><td rowspan="3">1992-2013</td><td rowspan="4">1985-2017</td></tr>
<tr><td>Oldott O</td><td></td></tr>
<tr><td>pH</td><td>-</td><td>1978-2006</td></tr>
<tr><td><u>Zala ÖP</u></td><td>mg l^{-1}</td><td></td><td></td></tr>
<tr><td><u>Wulka ill. Zala Q</u></td><td>$m^3 s^{-1}$</td><td rowspan="3">1992-2013</td><td rowspan="6"></td><td>1978-2006</td><td>1994-2010</td></tr>
<tr><td>Enterococcus</td><td rowspan="2">CFU 100 ml^{-1}</td><td rowspan="4"></td><td rowspan="9"></td></tr>
<tr><td>E-Coli</td></tr>
<tr><td>$\delta^{18}O$</td><td rowspan="2">ezrelék</td><td rowspan="3"></td><td rowspan="2">2000-2015</td><td>1970-1988</td></tr>
<tr><td>δ^2H</td><td rowspan="13"></td></tr>
<tr><td><u>Talaj P- és N-többlet</u></td><td>kg $ha^{-1} y^{-1}$</td><td>1978-2006</td><td rowspan="11"></td></tr>
<tr><td>Vezetőképesség</td><td>mS cm^{-1}</td><td rowspan="8">1992-2013</td><td>1985-2017</td><td rowspan="6"></td></tr>
<tr><td><u>Secchi-mélység</u></td><td>cm</td><td rowspan="5"></td></tr>
<tr><td><u>Szélsebesség</u></td><td>m s^{-1}</td></tr>
<tr><td><u>Napos órák száma</u></td><td>óra</td></tr>
<tr><td><u>Besugárzás</u></td><td>J cm^{-2}</td><td rowspan="2">1994-2010</td></tr>
<tr><td><u>Levegő-hőmérséklet</u></td><td rowspan="2">°C</td></tr>
<tr><td><u>Vízhőmérséklet</u></td><td>1985-2017</td><td>1978-2006</td><td></td></tr>
<tr><td><u>Csapadék</u></td><td>mm nap^{-1}</td><td rowspan="3"></td><td>1979-2006</td><td>1994-2010</td></tr>
<tr><td><u>Szennyvíz P</u></td><td>mg l^{-1}</td><td rowspan="2"></td><td>1978-2006</td><td></td></tr>
<tr><td><u>Felhőborítottság</u></td><td>tenths</td><td></td><td>1994-2010</td></tr>
<tr><td>Köppen-Gegier-klímazónakódok (Kottek et al. 2006)</td><td>-</td><td colspan="4" rowspan="2"></td><td rowspan="2"></td></tr>
<tr><td>Lat - Lon (EPSG: 3857) ELE</td><td>m</td><td></td></tr>
<tr><td colspan="2">Felhasznált program(ok)</td><td>R, IBM SPSS 26, Sufer 15, GS+ 10</td><td>R, IBM SPSS 26</td><td colspan="2">R</td><td colspan="2">R, Surfer 15, GS+ 10</td></tr>
</table>

3.2. Az alkalmazott módszertan rövid áttekintése

A bemutatott tanulmányokban vizsgált adatokat egy egységes rendszerben, négydimenziós térben lehet elképzelni (**3.2-1. ábra**), ahol a tengelyeket az idő (t), a vizsgált kémiai komponensek és a megfigyelési pontok adják. Itt a függőleges tengely (z) önmagában is több paramétert, több valószínűségi változót reprezentál(hat), ami az előbb mondottak szerint a paraméterek szintjén kibontva több dimenziónak felel meg, de - akárcsak a paraméterek halmaza - egyben kezelhető (Kovács, 2018). A (vízszintes) megfigyelési pontok (mf_n) tengelyén lévő pontokhoz térbeli koordináták (x_n; y_m) tartoznak, így egyszerűsíthető a négydimenziós adatstruktúra a háromdimenziós vizualizációként (Kovács et al., 2012c). Annak függvényében, hogy melyik tengelyt metsszük el egy arra merőleges síkkal (**3.2-1. ábra**: S1, S2 és S3), más-más vizsgálati módszer alkalmazható az adott síkban. Ezzel a lépéssel absztrakt módon rögzítjük egy időpontra, egy változóra, vagy egy térbeli mintavételi pontra a vizsgálatokat. Az alábbi beosztást a bemenő nyers adatok jellemzőinek ismeretében teszem meg, mivel az adatelőkészítés során számtalan transzformáció (pl. átlagolás a vizsgált időszakra stb.) végezhető, amellyel más és más beosztást kaphatunk.



3.2-1. ábra. Adathalmaz ábrázolása négy dimenzióban Kovács et al. (2012c) alapján**.**

Az alábbi bekezdések célja, hogy rövid betekintést adjanak, hogy a három síkban (**3.2-1. ábra**: S1, S2 és S3) milyen jellegű módszerek alkalmazhatók, kifejezett példákkal S2 és S3 sík elemeire, melyek közül több, e dolgozat gerincét is adja és melyek bővebb leírása a rövid értekezés különlenyomataiban találhatók. A felsorolt eljárások közül hármat részletesebben is bemutatok a következő 3.2.1 fejezetben, melyek úgy gondolom vagy jellegükben, vagy a dolgozatban bemutatott alkalmazásukban eltértek a szakterületen megszokottól.

S1 síkban jellemzően több mintavételi pontot érintő mintavételi kampányból származó, kvázi egy időben mért számos változó adatait figyelembe vevő vizsgálatokat lehet elvégezni, pl. Kovács et al. (2012c): 3.3. fejezet. Erre az esetre nem hozok példát a disszertációban, mivel a tárgyalt tanulmányokban ilyen nem fordult elő.

S2 síkban egy-egy változó idősorainak térbeli változékonyságát vizsgálhatjuk. Ez volt az eset a Fertő (**4.2-1. ábra**) fekáliás paramétereinek esetében, ahol az E. coli- és Enterococcus-csíraszámok idősorainak változékonyságát elemeztem főkomponensanalízissel (PCA; Hotelling, 1933; Tabachnick és Fidell, 2014), a tó mintavételi pontjainak együttes figyelembevételével, majd meghatároztam a külső magyarázó meteorológiai változók szerepét a főkomponensek alakulásában (Hatvani et al., 2018b). Továbbá, akárcsak a nyugat-antarktiszi jégfuratok esetében (**4.4-1. ábra**) vizsgált δ^{18}O-értékekre (Hatvani et al., 2017b), a Fertőn a két SFIB-ra is megnéztem a térbeli mintavételezési gyakoriság reprezentativitását a geostatisztika alapfüggvénye, a félvariogram (Matheron, 1963) segítségével. Az antarktiszi jégfuratokon végzett vizsgálatok végső terméke egy izotópos tájkép lett (Bowen, 2010; Bowen et al., 2019), amelyet regresszió-krigeléses (Hengl et al., 2007) interpolációval kaptam meg. Itt a regresszió segédváltozói a jégfuratok tengertől való távolsága és azok pozícióit leíró földrajzi változók (LAT, LON, ELE) voltak EPSG: 3857-es vetületben. A Mediterráneum δ_P-idősoraiból RMA regresszióval (IAEA, 1992; Harper, 2016) meghatározott LCsVV-k meredekségét és tengelymetszetét is egy-egy önálló paraméterként vizsgáltam, melyek térbeli változékonyságát három különböző interpolációs eljárással (IDW, RF, SVM) modelleztem; Hatvani et al. (2023): 2.3.3-as fejezet.

Végül S3 síkban egy-egy kiválasztott mintavételi helyen különböző változók idősorai kerülnek a vizsgálatok fókuszába. Ilyen eset volt jelen dolgozatban a KBVR-en a Zala befolyási pontján (**4.1-1A-B ábrák**: Zalaapáti mvp. (Z15)) az általános vízminőséget leíró változók vizsgálata (**3.1-1. táblázat**), amelyen dinamikus faktoranalízist alkalmaztam (DFA; (Geweke, 1977; Márkus et al., 1999; Ziermann és Michaletzky, 1995)), illetve a KBVR további élőhelyeit reprezentáló mintavételi pontokon az elsődleges tápanyagháztartás és a helyi meteorológiai viszonyok (**3.1-1. táblázat**) közötti periodikusságának elemzése wavelet-koherenciaanalízissel (WCA; Grinsted et al., 2004; Torrence és Compo, 1998). A Balaton esetében, akárcsak a KBVR-nél, több mintavételi pontot vettem figyelembe (egyet-egyet medencénként; **4.1-1B-C ábra**), tehát nem „tisztán" az S3 síkban vizsgálódunk, de a pontokon kapott eredmények közös értelmezésén kívül, azok kapcsolatait nem vizsgáltam statisztikai értelemben, ezért soroltam ezeket is S3-ba. A Balaton medencéiben PCA-val vizsgáltam meg, hogy adott ponton számos vízminőséget jellemző változó milyen közös lineáris változékonyságot mutat, és ez mennyire köthető a Zalából érkező összes foszfor éves terheléseihez (**3.1-1. táblázat**).

Az említett, a kutatásban alkalmazott eljárások részletes leírása és a hozzájuk kapcsolódó módszertani megközelítések, a balatoni PCA- és trendvizsgálatok esetében Hatvani et al. (2020a), a KVBR-en végzett DFA- és WCA-vizsgálatok vonatkozásában rendre Hatvani et al. (2015) és Hatvani et al. (2017a), a Fertő fekáliás és vízminőségi paraméterein végzett PCA- és geostatisztikai vizsgálatoknál Hatvani et al. (2018b), míg a Mediterráneum LCsVV-inek meredekség- és tengelymetszetéből és az antarktiszi jégfuratok δ^{18}O-összetételéből modellezett izotópos tájképek létrehozásának és leírásának vonatkozásában pedig rendre Hatvani et al. (2023) és Hatvani et al. (2017b) publikációkban találhatja meg az olvasó.

3.2.1. Néhány alkalmazott módszer részletesebb ismertetése

A következőkben egy-egy bekezdésben összefoglalom a kombinált klaszter- és diszkriminanciaanalízis (CCDA; kollégákkal közösen fejlesztett módszer), a főkomponens-analízis (PCA) és a DFA lényegét, végül kitérek egy bekezdésben a Mediterráneum LCsVV-paramétereire alkalmazott becslő eljárásokra.

A CCDA (Kovács et al., 2014a; Kovács et al., 2014b) két ismert módszer kombinációjából áll, a hierarchikus klaszter- (HCA; Day és Edelsbrunner, 1984; Driver és Kroeber, 1932) és a lineáris diszkriminanciaanalízisekből (LDA; Duda et al., 2012; Fisher, 1936). Előbbi a csoportok hierarchiáját hozza létre, amelyben a legelemibb szinten minden mintavételi hely saját csoportot alkot, míg a legdurvább felosztásban minden mintavételi hely azonos csoporthoz tartozik egy adott transzformált távolság függvényében. Az LDA pedig olyan síkokat határoz meg, amelyek az adott csoportokat optimálisan választják el. Eredményként a síkok által helyesen klasszifikált megfigyelések százalékát kapjuk (Kovács, 2018). A két módszer összevonásának esszenciája a hagyományos csoportosító eljárásokkal ellentétben, hogy objektív mérőszámot rendel a HCA eredményéhez, így segítségével homogén és optimális csoportok is meghatározhatók (Kovács, 2018).

PCA a hidrológia területén is egy régóta alkalmazott dimenziócsökkentő eljárás; pl. Snyder (1962); Wallis (1965). Lényege, hogy a vizsgált változók lineáris transzformációjával új ortogonális főkomponenseket hoz létre, amelyeket az egymással korreláló eredeti változók határoznak meg. Ezen főkomponensek száma azonos az eredeti változók számával, és az eredeti teljes variancia egy adott százalékát magyarázzák, rendre szigorúan monoton csökkenő formában (Tabachnick és Fidell, 2014). Noha van példa a PCA hidrológiai alkalmazásában, hogy a lehetséges külső magyarázó tényezőket időben (Çamdevýren et al., 2005; Hatvani et al., 2018b; Page et al., 2012) vagy térben (Olsen et al., 2012) végzett PCA-eredményekkel (PC-idősorokkal) korreláltatva keresik, de nem gyakori. A Fertőn végzett elemzések más megközelítést alkalmaztak. A PCA balatoni alkalmazásával ellentétben, ahol mintavételi pontonként készült el az ott mért több változó idősorából a korrelációs mátrix, amely a PCA bemeneti információját adta, a Fertőn az E. coli- és Enterococcus-csíraszám különböző mintavételi pontokon mért idősoraiból készült egy-egy korrelációs mátrix, és ezekből a PCA. Az így kapott térbeli főkomponens-„adatsorokat" korreláltattam a rendelkezésre álló lehetséges magyarázó változókkal (pl. szélsebesség, csapadék, napos órák száma, stb.; **3.1-1. táblázat**), így megkapva a fekáliás forrópontokat (Hatvani et al., 2018b). A PCA alkalmazása előtt meggyőződtem róla, hogy mind a Balaton, mind pedig a Fertő esetében a bemenő adatok autokorreláltsága relatíve gyenge (lag-1 r < 0,7), így nem volt indokolt DFA alkalmazása.

A DFA a klasszikus faktormodellezés alapgondolatait az idősorok empirikus struktúrájának figyelembevételével alkalmazza, ötvözve az idősor vektorok és mátrixok dinamikus tulajdonságainak vizsgálatában alkalmazott idősorelemzési eljárásokat a faktoranalízis információsűrítő dimenziócsökkentő eljárásaival (Ziermann és Michaletzky, 1995). Azonban abban különbözik más eljárásoktól (pl. faktoranalízis, vagy PCA), hogy a dinamikusfaktor-modell a faktoridősorokat autoregressziós jellegűnek tekinti, ezáltal kezeli idősorok esetében az autokorreláltság problémáját és egyben figyelembe veszi a komponensek közötti késleltetett összefüggéseket, amire a PCA nem alkalmas (Geweke, 1977; Ziermann és Michaletzky, 1995). Továbbá – akár csak a PCA – alkalmas, hogy a kapott faktorokból az eredeti idősorok reprodukálhatók legyenek. A kapott faktoridősorokat később független változók idősoraival korreláltatva lehet háttértényezőket azonosítani, pl. Hatvani et al. (2015); Kisekka et al. (2013); Kovács et al. (2015b); Magyar et al. (2021); Muñoz-Carpena et al. (2005), és a hidrogeológiában először Kovács et al. (2004).

A Mediterráneum LCsVV-tengelymetszetének és meredekségének térbeli változékonyságát három predikciós technikával határoztam meg 2000 és 2015 közötti időszakra. Az alapvetően hagyományos interpolációs eljárás, a távolsággal fordítottan arányos súlyozás (IDW) (Webster és Oliver, 2008), valamint két geomatematikai gépi tanulási (ML) módszer: véletlen erdő (RF) (Li et al., 2011; Liu et al., 2012) és az ún. „support vector machine" (SVM) (Cortes és Vapnik, 1995) eljárást alkalmaztam. A kutatás különlegessége, hogy míg az IDW esetében lineáris ($\alpha = 1$) és nemlineáris ($\alpha = 2$) súlyozást is vizsgáltam, addig az ML eljárásoknál a lehetséges prediktorok hat kombinációjával történt a térbeli becslés, amelyek befolyásolhatják a csapadék stabilizotóp-összetételét, következésképpen az LCsVV-k meredekségét és tengelymets zetét. Ezek (i) LAT, LON és a Köppen–Geiger-klímazónák kódjai, (ii) LAT, LON, és ELE Liotta et al. (2006) alapján, (iii) LAT, LON, ELE, és a Köppen–Geiger-klímazónák kódjai, (iv) az állomások távolságmátrixa és a Köppen–Geiger-klímazónák kódjai kódjai Hengl et al. (2004) alapján, (v) a távolságmátrix és ELE, végül (vi) a távolságmátrix, ELE és a Köppen–Geiger-kódok voltak. A modellek teljesítményének értékelése két validáló adathalmaz segítségével történt, elsősorban a négyzetes hibák mediánjainak figyelembevételével; részletek Hatvani et al. (2023): 2.3.4. fejezetében olvashatók.

4. EREDMÉNYEK ÉS DISZKUSSZIÓ

A doktori értekezés keretében vizsgált víztesteket és makrorégiókat jellemző vízminőségi, izotóphidrológiai és egyéb környezeti adatokon végzett geomatematikai elemzésekből kapott eredmények közül először egy kis vízhozamú vízfolyás–eutróf tó–wetland–sekély tó kaszkádrendszer (rendre Zala – Hídvégi-tó – Ingói- és Zimányi-berkek – Balaton) (4.1. fejezet) és a Fertő elemzéséből kapott megállapításokat taglalom (4.2. fejezet), majd a δ_P térbeli becsléséből kapott eredményeket mutatom be (4.3. és 4.4-es fejezetek). A 'rövid értekezés' jellegéhez alkalmazkodva, először a mintaterületeket és a kapcsolódó konkrét tudományos kérdéseket, majd a vonatkozó eredményeket, és az ezekből levont *tézisértékű megállapításokat* mutatom be. Szintén a 'rövid értekezés' jellegéből adódóan, az irodalmi hivatkozásoknál elsősorban a dolgozat gerincét adó saját tanulmányokra helyezem a hangsúlyt, amelyekben azonban részletesen ki van fejtve egy-egy probléma és meghivatkozva a legfontosabb vonatkozó irodalom.

4.1. Egy kis vízhozamú vízfolyás (Zala) – eutróf tó (Hídvégi-tó) – vizes élőhely (Ingói- és Zimányi-berkek) – sekély tó (Balaton) kaszkádrendszer

4.1.1. A Balaton és a Kis-Balaton Vízvédelmi Rendszer (KBVR) bemutatása

A Balaton Közép-Európa legnagyobb sekély tava (vízfelszíne 596 km^2), átlagos vízmélysége 3.2 m) (**4.1-1C ábra**). A Zala folyó adja a Balaton természetes vízutánpótlásának kb. 50%-át és tápanyag-utánpótlásának 35-40%-át (Istvánovics et al., 2007). A Balaton legnyugatibb és egyben legkisebb medencéjébe, a Keszthelyi-medencébe torkollik. Legjelentősebb kifolyója a Sió-zsilip és -csatorna, amelyet vízszintszabályozási céllal építettek a XIX. században.

A XX. század második felében a Balaton vízgyűjtőjében megnövekedett az antropogén aktivitás: a népességnövekedés, az üdülés, a szennyvízkibocsátás, a műtrágyahasználat megemelte a tóba jutó tápanyagok mennyiségét (Hatvani et al., 2015; Herodek, 1984; Herodek et al., 1982; Somlyódy et al., 1983), és vízminőségromlást eredményezett (Sebestyén et al., 2017). Az 1980-as évek végére a Zala által szállított P mennyisége megduplázódott az előző évtized elejéhez képest (Istvánovics et al., 2007; Sagehashi et al., 2001).

E negatív folyamatok megakadályozására a hatóságok egy akciótervet hoztak létre az alábbi meghatározó elemekkel:

(i) regionális szennyvízelvezetés kiépítése a tó körül, szennyvizek kivezetése a vízgyűjtőről;
(ii) a tó nyugati vízgyűjtőjén új szennyvíztisztító telepek létrehozása és tápanyag-eltávolítás bevezetése;
(iii) élőállat-tenyésztéssel foglalkozó telepek szabályozása (hígtrágyás tartás tiltása); és
(iv) a Kis-Balaton Vízvédelmi Rendszer (KBVR) létrehozása azzal a céllal, hogy visszaállítsák a Zala-völgy 1860-as éveket megelőző időszakra jellemző hidrológiai viszonyait, ezzel megnövelve a víz tartózkodási idejét a Keszthelyi-medence előtt a

hajdani Kis-Balaton területén, így visszatartva azokat a Balatontól (Hatvani et al., 2020a; Somlyódy és van Straten, 1986; Tátrai et al., 2000).

A KBVR-t az egykori Zala-torkolatban lévő vizes élőhely, wetland helyén hozták létre (Lotz, 1988), mert annak zsugorodásával csökkent szűrőképessége, így a Zala szinte közvetlenül szállította a diffúz és pontforrás eredetű tápanyagokat a Keszthelyi-öbölbe. A KBVR 2015-ig két fő működési egységből állt (**4.1-1B ábra**). Egyik a Hídvégi-tó (területe 18 km^2, a víz tartózkodási ideje kb. 30 nap), amelyet ötlépcsős elárasztás után 1985 nyarán üzemeltek be. Az eredeti tervek szerint egy nádassal benőtt élőhely lett volna azzal a céllal, hogy visszatartsa a szervetlen növényi tápanyagokat a Balatontól; ezzel szemben egy nyílt vizű eutróf tóvá alakult (Korponai et al., 2010) (**4.1-1B ábra: Z11 pont**). A második fő egység a Fenéki-tó, amelynek az eredeti tervek szerint makrofita vegetációval dúsan borított 51 km^2-e ~90 nap alatt engedte volna a Zala vizét a Balatonba. 1992-ben még csak az Ingói-berek területén fekvő 16 km^2-es területet tudták bevonni a rendszerbe, hogy a Hídvégi-tóból elfolyó, algában gazdag víz itt tisztuljon tovább, valamint, hogy a több száz évig pangóvizes terület átfolyásos rendszerré alakításának negatív hatásai minél hamarabb lejátszódjanak. Az Ingói-berek tovább „osztható" egy zavartalan (**4.1-1B ábra: Kb210 pont**), és egy zavart vizes élőhelyre (**4.1-1B ábra: Z27 pont**). Utóbbiba a Zala vízhozamához képest 40%-os víztöbblet érkezik különböző csatornákon és a Fenéki-tavon keresztül (Hatvani et al., 2014a), ez található a KBVR Balatonhoz közelebb eső felén. A gazdasági környezet változásának hatására azonban a KBVR kivitelezése leállt, és a munkák évekig csak a meglévő és építés alatt álló műtárgyak és egyéb létesítmények karbantartására korlátozódtak. A munkálatok végül teljes egészében 2015-ben fejeződtek be. A Zalavári belvízöblözet elárasztásával, rugalmas vízkormányzás biztosítása mellett, részlegesen leválasztották az Ingói-berket a Fenéki-tó többi részéről. Ezzel kívánják az egykori Kis-Balaton Wetland (vizes élőhely) vízjárását és természetes szűrőszerepét pótolni a Balaton vízminőség-védelmének érdekében, miközben az Ingói-berek természeti értékeit is meg lehet óvni.

A fent bemutatott intézkedésekkel és a műtrágyahasználat tizedére csökkenésével az 1990-es évekre nagyságrendileg felére csökkent a Balatonban érkező összesfoszfor-terhelés mértéke az 1980-as évekhez képest, így az eutrofizációs folyamatok visszafordultak (Hatvani et al., 2014a; Istvánovics et al., 2007). A Balaton esete nemzetközi szinten is „sikertörténet" lett az eutrofizáció kezelésének kérdésében (Ho et al., 2019).

Összeségében a Zala–KBVR–Balaton egység egy olyan kaszkádrendszernek tekinthető (Hatvani et al., 2014a, 2017a; Honti et al., 2020), amelynek elemei egy alacsony vízhozamú vízfolyás, egy eutróf tó, egy vizes élőhely (zavartalan és zavart 2015-ig), valamint egy sekély tó (Balaton). Az élőhelyek között ökoton (Gosz, 1993) csak minimálisan van jelen (Hatvani et al., 2017a), az elsősorban a wetland és annak partvonala mentén található (Zlinszky et al., 2012). Az átmeneti élőhelyek „hiánya" miatt a rendszer elemei így valóban kezelhetők jelen vizsgálatok szempontjából négy „fekete dobozként". Jelen értekezésben is ezt a logikát követve mutatom be az egymáshoz szorosan kapcsolódó eredményeket, amelyek egy-egy külön tanulmányban jelentek meg.

A Balaton- és a KBVR vízminőségéhez kapcsolódó kutatási célok

Ahhoz, hogy minél hatékonyabban védekezhessünk a Balaton eutrofizálódása ellen, szükséges részleteikben megérteni a Zala által szállított növényi tápanyagok eredetét és ezek mennyiségét irányító háttérfolyamatokat. Ennek érdekében dinamikus faktoranalízissel (DFA; Geweke (1977)) megvizsgáltam, hogy milyen külső diffúz vagy pontforrás eredetű magyarázó tényezők (pl. trágyafelhasználás, a Zalába vezetett tisztított szennyvíz stb.)

befolyásolják a Zala hosszú távú (1978–2006) vízminőségi változásait, és ezek hogyan változnak időben (Hatvani et al., 2015; Sisák, 1993).

Továbbá, mivel a Zala folyóból érkező tápanyagok a KBVR élőhelyein keresztül „szűrve”, visszatartva jutnak a Balatonba, érdemes megvizsgálni, hogy ezt az alacsony vízhozamú vízfolyás – eutróf tó – vizes élőhely alkotta rendszert jellemző vízminőségi változók periodikus viselkedése mennyire képes követni az elsősorban besugárzás által vezérelt ún. hidro(geo)kémiai évszakosságot (Kolander és Tylkowski, 2008; Tanos et al., 2015). Amennyiben fel lehet tárni az okokat, hogy miért képes vagy nem pl. az eutróf előhely (Hídvégi-tó) vagy a vizes élőhely (ennek is különböző egységei) jobban, vagy kevésbé követni a meteorológia által vezérelt évszakosságot, az egy új betekintést ad az élőhelyek folyamataiba, így célozottabb intézkedéseket tehet a NYUDUVIZIG azok megóvására, illetve még hatékonyabb üzemeltetésére a Balaton vízminőség-védelme céljából (Hatvani et al., 2017a).

Végül átfogó képet adok a Balaton trofikus állapotának elmúlt 30 évben bekövetkezett alakulásáról, hiszen Istvánovics et al. (2007) meghatározó munkája óta nem született hasonló tanulmány. A kifejezett cél, hogy (i) három évtized távlatából statisztikailag megvizsgáljam, elkülöníthetők-e különböző vízminőségi/trofikus állapotokkal jellemezhető időszakok a tó történetében (ii), és amennyiben igen, meghatározhatók-e ezekben robusztus, hosszú távú trendek a trofikus paraméterekre vonatkozóan, valamint, hogy (iii) ezek mennyiben változtak időben és a Balaton medencéiben térben. Továbbá (iv) cél annak meghatározása is, hogy a vizsgált több mint 30 év alatt a Zalából érkező külső összesfoszfor-terhelések mennyiben befolyásolják a medencék vízminőségét (Hatvani et al., 2020a).

4.1.2. A Balaton és a KBVR vízminőségének és trofikus állapotának változása térben és időben

A Zala–KBVR–Balaton kaszkádrendszerbe a Zala folyón keresztül érkezik a tápanyagok meghatározó hányada (**4.1-1A-B ábrák**). Ahhoz, hogy a leghatékonyabban lehessen védekezni az érkező tápanyagtöbblet ellen az oligotróf állapotok eléréséhez, a Hatvani (2014): 5.2.2. fejezetében bemutatott, majd tovább gondolt, és mind módszertanában, mind értelmezésében kibővített kutatás során (Hatvani et al., 2015) elsőként 21 ún. „response-idősorból” (**3.1-1. táblázat**) meghatároztam a dinamikus faktoridősorokat/közös trendeket az 1978–2006-os időszakra (21 db). Következő lépésként AIC, BIC és RMSE metrikák segítségével ezek közül az első három faktort választottam ki (Hatvani et al. (2015): 2. táblázat) , és megvizsgáltam, hogy az 1991–’92-ben a „response-változók” varianciájában tapasztalható egyértelmű változás (Hatvani et al. (2015): 2a ábra) megjelenik-e párhuzamosan a háttértényezőkben.

A két időszakra (1978–1991 és 1992–2006) is elvégeztem külön-külön a DFA-t. Meghatároztam három-három első faktor paramétereinek súlyát a megfigyelési pontokon, végül a faktoridősorok külső, magyarázó idősorokkal való kapcsolatát. Az 1978–1991-es időszakban (Hatvani et al. (2015): 4. ábra) az első faktor az éves csapadékmennyiség-idősorral korrelált jól ($r(13) = ,77$, $p = ,002$), míg a második a Zala-vízgyűjtő talaja P-többletének idősorával ($r(14) = ,67$, $p = ,0012$) és a szennyvíz P-tartalmával (WW-P; $r(14) = ,59$, $p = ,0012$). 1992–2006-ban az első faktor a Zala vízhozamával és a WW-P-vel mutatott szignifikáns lineáris kapcsolatot ($r(15) = ,74$ és 0,65, $p = ,000$), míg a talaj P-többlettartalmával nem volt meghatározó kapcsolat ($r(15) = -,076$, $p = ,000$) (Hatvani et al. (2015): 5. ábra). A második faktorhoz nem lehetett magyarázó háttértényezőt találni, de a harmadik faktor a vízhőmérséklettel mutatott már értelmezhető szignifikáns kapcsolatot ($r(15) = ,77$, $p = ,000$). A paraméterek súlyából ítélve azt feltételezem a második faktor esetében, hogy azt a talaj szervesanyag-tartalmát és redox állapotát leíró független változó magyarázhatta volna meg.

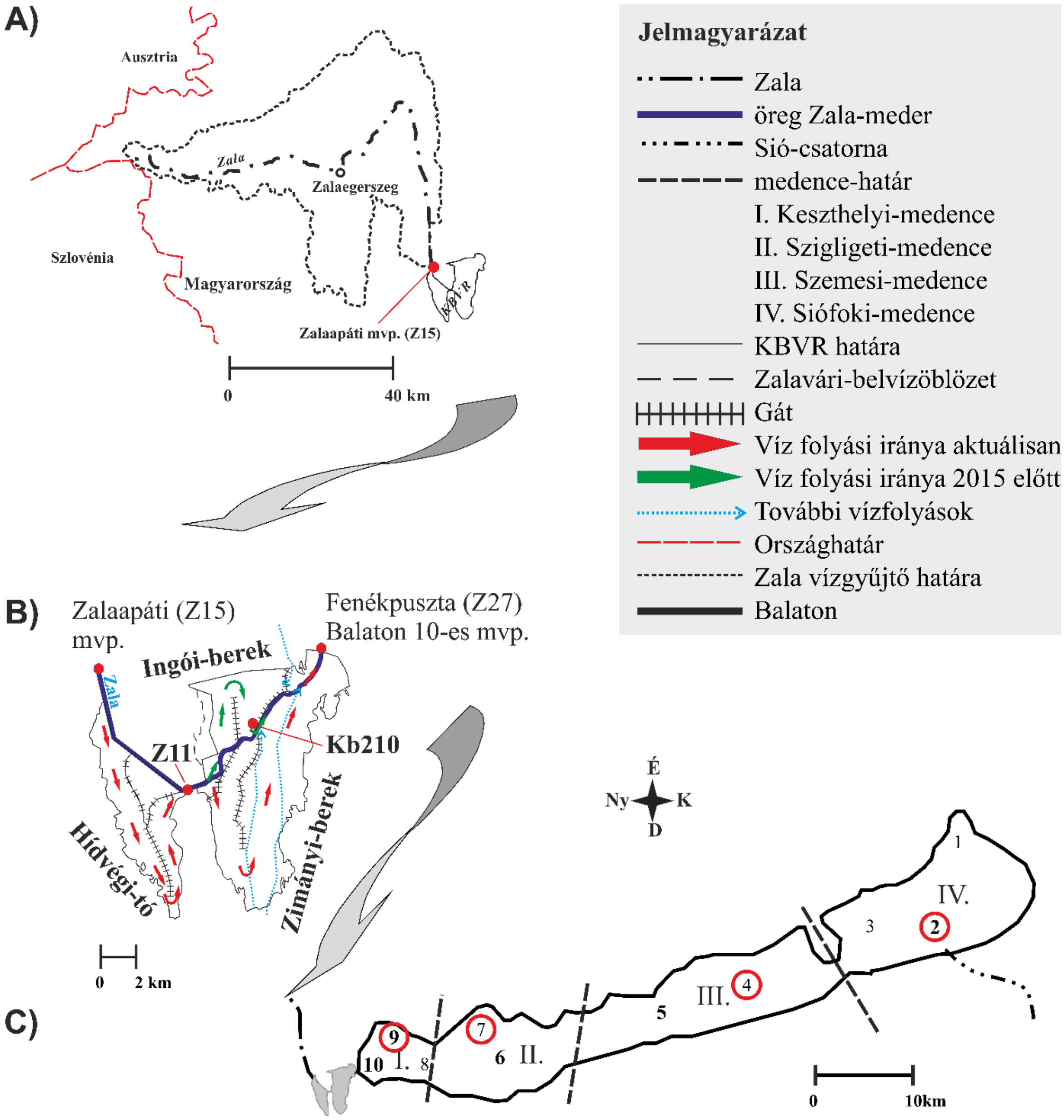
Összességében *dinamikus faktoranalízissel időben változó intenzitású és eltolt korrelációs struktúrájú háttértényezőket határoztam meg egy kis vízhozamú vízfolyás–eutróf tó–wetland–sekély tó fő bemeneti pontján. Kimutattam, hogy e kaszkádrendszer bemeneti oldalán a diffúz mezőgazdasági terheléshez és jellemzően pontforrás eredetű szennyvízhez köthető folyamatok szerepét átvették a természeti folyamatokat inkább jellemző háttértényezők (pl. vízhőmérséklet). Megállapítottam, hogy a tapasztalt jelenség hátterében elsősorban a vízminőség-védelmi intézkedések és az antropogén tevékenység csökkenése, az 1990-es években bekövetkezett nagymértékű mezőgazdasági átrendeződés állhat* (Hatvani et al., 2015). *Módszertani szemszögből megállapítottam, hogy a Zala folyó vízgyűjtőjéről érkező terhelések és azok megjelenése a KBVR-en, beleértve az előbbiekben említett vízgyűjtőre ható külső tényezőket, időben eltolva jelennek meg, és az ilyen folyamatok vizsgálata a dinamikus faktoranalízist követelik meg, mivel ez képes figyelembe venni az eltolt korrelációs struktúrákat, amelyek a vizsgálat tárgyát képezték.*

Wavelet-koherenciaanalízissel (WCA) megvizsgáltam a Zala és a KBVR élőhelyein (4.1.1. fejezet), hogy az ott mért NO_3-N, ÖN, ÖP, ORP és Ö Leba. mennyiben képes követni a magyarázó meteorológiai változók (**3.1-1. táblázat**) periodikusságát, illetve milyen időbeli eltolódás tapasztalható ezek periodikus jelei között a 4.1.2-es fejezetben megfogalmazott hipotézis nyomán. A kapott eredmények kiértékelésénél a WCA három alábbi tulajdonságát vettem figyelembe:

- jelen van-e koherens periódus a vízminőségi és meteorológiai változók között, és ha igen, akkor a vizsgált időszak hány százalékában;
- mekkora a maximális „Global Wavelet Power” (milyen erős a kapcsolat); és
- mekkora a fáziseltolódás az esetlegesen feltárt koherens periódusok között, amely esetében a vízminőségi változóknak kell időeltolódással követniük a független változók jeleit.

Csak azon jeleket vettem figyelembe, amelyek szignifikánsak voltak (α = ,01) 1 000 db AR(1) idősorral szemben; lásd Torrence és Compo (1998). Az eredmények részletesen Hatvani et al. (2017a)-ban olvashatók, az alábbiakban csak a legfontosabb és átfogó eredményeket emelem ki.

Az elemzés során a wetland élőhelyet csak az Ingói-berek képviselte, mivel a Zimányi-berek ekkor még csak közel egy éve került elárasztásra (**4.1.-1B ábra**: víz folyási iránya 2015 előtt). Az Ingói-berek továbbá két külön térrészre különült el, egy ún. zavartalan, makrofita dominálta wetlandre (Kb210 pont felett) és a wetland azon részére, amire délről további vízfolyások érkeznek (**4.1.-1B ábra)**; utóbbit „zavart wetlandnek” nevezem (Hatvani et al. (2017): 2.1. fejezet). A kaszkádrendszer élőhelyei közül az eutróf tó (átlagosan 80%-ban) és a zavartalan wetland (79%-ban) követte a vizsgált 17 éves időtartamban a meteorológiai változók (**3.1-1. táblázat**) periodikus viselkedését, míg a Zala és a „zavart wetland” már kevésbé (rendre 74% és 63%-ban). Ez magyarázható a vízmennyiség kb. 40%-os többletével, amely éven belül egyenetlenül elosztva (Hatvani et al., 2017a; Hatvani et al., 2014a) változó oldott tápanyagmennyiséggel érkezik a „zavart wetland” területére, ezáltal azt éves periódus szempontjából a Zala vízminőségi változóinak viselkedéséhez teszi hasonlóvá (Hatvani et al., 2017a) (**4.1-1B ábra**).



4.1.-1. ábra. A Zala és vízgyűjtője A) - Kis-Balaton Vízvédelmi Rendszer B) – Balaton C) kaszkádrendszer. A piros pontok (A és B panelek) jelölik a KBVR mintavételi pontjait: Zalaapáti (Z15), Z11 és Fenékpuszta (Z27), ami egyben a Balaton legnyugatibb (10-es) mintavételi pontja is (C panel). A piros körök a Balaton 10 mintavételi pontja közül azokat jelölik, melyek 2005 után folyamatosan működtek és Hatvani et al. (2020a) munkájában vizsgálva voltak. A Balaton mintavételi pontjai rendre: Balatonfűzfő (1), Siófok-Alsóörs (2), Balatonfüred (3), Balatonakali (4), Révfülöp (5), Szigliget (6), Szigligeti-öböl (7), Balatongyörök (8), Keszthelyi-öböl (9), Zala- torkolat (10; Z27).

A magyarázó és vízminőségi változók közötti fáziseltolódások legfontosabb üzenete, hogy az eutróf tóban az ÖP és Ö Leba egy-két hónap késéssel mutatja az éves periódus csúcsát, mint a besugárzás, és két-három hónap eltolódással a levegőhőmérséklethez képest. Eközben az ORP esetében nem mutatható ki értelmezhető fáziseltolódás az eutróf tóban. Ez a várakozásoknak megfelelő eredmény, és azt tükrözi, hogy az eutróf tóban a lebegőanyag-tartalom főként algából áll, és az ORP-nak csak egy alacsonyabb amplitúdójú éves ciklicitása jelenik meg (Hatvani et al. (2017a): A3b ábra).

Míg a NO_3-N az egész rendszerben antifázisban van a besugárzással és levegőhőmérséklettel, addig az ÖN a Zalában antifázisban, az eutróf tóban pedig fázisban volt, majd a wetland területén nem volt már értelmezhető az ÖN fáziseltolódása. Ez a mintázat azzal magyarázható, hogy a denitrifikáció nyáron aktívabb a rendszerben, ezért az NO_3-N mennyisége nyáron lecsökken, elsősorban a KBVR-ben (Hatvani et al. (2017a): 6. ábra), ami a Global Wavelet Powerben is megjelenik (Hatvani et al. (2017a): 3. táblázat). Eközben az ÖN nem mutatott egyértelmű fázist a besugárzással a Zalában, feltehetőleg a szerves N-tartalom változása miatt (Hatvani et al. (2017a): 4.2.2. fejezet), majd az eutróf tóban a NO_3-N-nel ellentétes pozitív fázist mutatott a besugárzással, mivel a szervetlen N-felvétel nyári fokozódásának hatására a nitrátot felveszik az algák (Reay et al., 1999), de az ÖN nyári csúcsa megmarad (Hatvani et al. (2017a): 6. ábra). A KBW-en pedig az ott zajló összetett és sokszor évszakos szinten is ellentétes irányú folyamatok miatt (algák nyári lebomlása, N-felvétel a magasabb rendű növényzet által; lásd Hatvani et al. (2017a): 4.2.2-es fejezet) nem jelenik meg egyértelmű fázis vagy antifázis a meteorológiai független változók és az ÖN között.

A fenti eredmények kiemelten fontos információt szolgáltatnak a wetland sérülékenységére vonatkozóan, a KBVR 2015-ös átalakítása után is. Ehhez kapcsolódóan 2020-ban a NYUDUVIZIG-gel közösen megkezdődött egy izotóphidrológiai monitoring, melynek célja, hogy feltárja a KBVR különböző élőhelyire érkező felszíni és lehetséges felszín alatti hozzáfolyások hatását és arányát.

Összefoglalva, *egy kis vízhozamú vízfolyás–eutróf tó–„zavartalan wetland”–„zavart wetland” alkotta rendszer különböző élőhelyein wavelet-koherenciaanalízissel kimutattam, hogy az elsődleges tápanyagformák jelentősen eltérő mértékben mutatnak hidro(geo)kémiai évszakosságot. Az eutróf tóban zajló folyamatok voltak képesek a legjelentősebb mértékben tükrözni a meteorológiai változók évszakosságát. Az ún. „zavart wetland” pedig a Zalához hasonló viselkedést mutat, mert a Marót-völgyi-csatorna, a Zala-Somogyi-csatorna és az Egyesített-övcsatorna 40%-nyi vízhozamtöbblete rendszertelenül érkezett az Ingói-berekbe* (Hatvani et al., 2017a).[4]

A kaszkádrendszer utolsó eleme a Balaton (**4.1-1C ábra**). A tó történetében kimagaslóan fontos szerepet töltött be az eutrofizáció, ami számtalan esetben kritikus mértéket öltött (Herodek, 1984; Istvánovics et al., 2007; Istvánovics et al., 2022; Pomogyi, 1996; Somlyódy és van Straten, 1986). A vizsgálatok során a Balaton négy medencéjéből egy-egy mintavételi pontról három évtizednyi (1985–2017) adatot vizsgáltam meg kollégákkal és szakdolgozó hallgatómmal (**4.1-1C ábra)**. A vizsgálatok fókuszában az egyik legelterjedtebben alkalmazott trofitásindexnek (Istvánovics, 2009), a Nemzetközi Gazdasági és Fejlesztési Szervezet (OECD) trofitásindexének (Vollenweider és Kerekes, 1982) fő paraméterei álltak (ÖP és Chl-a), amelyek az újabb indexek alapjait is képezik (Markad et al., 2019; Wen et al., 2019).

A Balatonban mért és a vizsgálatba bevonható számú adattal rendelkező vízminőségi változók alapján (**3.1-1. táblázat**) kombinált klaszter- és diszkriminanciaanalízissel három vízminőségi időszakot különítettem el a tó történetében (1985–1994, 1995–2003 2004–2017; (Hatvani et al. (2020a): 3. ábra), majd ezen időszakokban Sen’s slope-vizsgálattal (Sen, 1968) meghatároztam az ÖP, ORP és Chl-a hosszú távú trendjét medencénként és időszakonként. Utolsó lépésként PCA-val kapott medencénkénti főkomponenseket és a Zalából érkező éves ÖP-terhelést autokorrelációra korrigálva (Macias-Fauria et al., 2012) korreláltattam, hogy megvizsgáljam, mennyiben lehet meghatározó a külső terhelés a medencéket leíró vízminőségi változók varianciájának alakulásában (Hatvani et al. (2020a):

[4] A tézisértékű megállapításokat a dolgozat vonatkozó fejezeteiben *dőlttel szedtem*.

2.2.2. fejezet). Más szóval, azt kívántam kideríteni, hogy a Balaton egyik fő limitáló tényezője (a foszfor (Herodek, 1984)) mennyiben befolyásolja a vízminőséget a tó tengelye mentén, ahogy távolodunk annak fő befolyási pontjától, a Keszthelyi-medencétől, ahová a Zala vize érkezik (**4.1-1C ábra**). Az eredmények alapján megállapítható, hogy a vizsgált 33 évben a trofitást leíró paraméterek koncentrációja időben először a Zala befolyásánál a nyugati medencé(k)ben kezdett csökkenni az 1990-es évek derekán (a 4.1.1-es fejezetben taglalt intézkedéseknek és eseményeknek köszönhetően), majd terjedt kelet felé (Hatvani et al. (2020a): 2. ábra és 2. táblázat). Egyúttal megerősítettem a már korábbról is ismert tényt (Herodek, 1984; Istvánovics et al., 2007), hogy a külső ÖP-terhelés egyre kevésbé befolyásolta a keletibb medencék vízminőségét, a szervetlen tápanyagok (pl. P-formák) egyre kisebb mértékben határozták meg a vízminőségi változók varianciáját az egyes medencékben (Hatvani et al. (2020a): 3. táblázat).

A Balaton trofitási fokának csökkenése az intézkedések ellenére, azonban csak időben eltolva indult meg, többek között az üledékből való P-visszaoldódás következményeként (Istvánovics et al., 2004). Azon években, amikor a belső P-terhelés évtizedes maximumait éri el, erős korreláció tapasztalható a fitoplankton-biomassza és a becsült mobil-P-koncentrációk között (Istvánovics et al., 2022), amit az üledék karbonáttartalma is jelentősen befolyásol (Istvánovics, 1988). A tavak időben eltolt válasza a külső terhelés csökkentésére megnehezíti a jelenség vizsgálatát (Sas, 1990), és sok esetben ellehetetleníti, hogy évtizedesnél kisebb időtartamot vizsgálva korrelációt találjunk a külső terhelés csökkentése és a vízminőség javulása között.

Több mint harmincévnyi, vízminőséget és egyben trofikus állapotokat leíró adat új megközelítésű elemzése nyomán az eddigi irodalommal egybecsengő kép rajzolódik ki a Balaton trofikus állapotát illetően. Eddig fel nem dolgozott adatokat is elemezve megállapítottam, hogy (i) *a Balaton trofikus állapota a nyugati medencék hipertróf állapotából a Siófoki-medencében már oligotróf állapotokat is mutatott klorofill-a alapján. Ez a javuló trend pedig az egész tóra jellemző az 1990-es évek közepétől kezdődően, amely azonban időben késleltetve jelent meg a medencékben kelet felé haladva. Ezt megerősítette a vizsgált időszakok szétválása és a Sen-féle hosszú távú trendvizsgálat is. Továbbá, (ii) a vízminőségi változókon végzett főkomponens-analízis faktoraiban kelet felé csökkent a tápanyagok szerepe, és ezzel párhuzamosan gyengült a főkomponensek korrelációja a Zalából érkező ÖP-terheléssel, azt jelezve, hogy a medencék térben is máshogyan reagáltak a terhelésekre (Hatvani et al., 2020a).*

Az eredmények megerősítik, hogy a Balaton trofikus állapotának javulása nagymértékben függött (i) a tó hidromorfológiai jellemzőitől, (ii) a külső terhelés csökkentésére irányuló intézkedésektől (pl. szennyvíz P-tartalmának csökkentése, csatornázás), amit alátámaszt a Zalából érkező, szignifikánsan csökkenő ÖP-terhelés, (iii) a rendszerváltás hatására „megroppanó” ipartól és elsősorban mezőgazdaságtól, (iv) a meteorológiai állapotoktól, amelyek számtalan esetben meghatározták a P-reszuszpenziót az üledékből, valamint az ORP deszorpcióját. Az eredményeim, összevetve nemzetközi esettanulmányokkal (Hatvani et al. (2020a): 4.3-as fejezet), kihangsúlyozzák, hogy csak átfogó külső tápanyagterhelés-csökkentéssel lehet sikeresen visszafordítani hasonló sekély tavak eutrofizálódását, amelynek szűk keresztmetszete a bejutó P-mennyiség csökkentése (Hatvani et al., 2020a).

4.2. Európa legnyugatibb sekély szikes tavának fekáliás szennyeződése

4.2.1. A Fertő bemutatása – fókuszban az osztrák tórész

A Fertő a legnyugatibb szikes tó Európában, és egyben Ausztria legnagyobb sekély tava. Vízfelülete 309 km^2, átlagos mélysége pedig 1,1 m, területének kb. negyede található Magyarországon (75 km^2) (Dinka et al., 2016). A tó Ramsari védelem alatt áll, és Magyarország, ill. Ausztria Nemzetközi Nemzeti Parkja, Egyesült Nemzetek Nevelésügyi, Tudományos és Kulturális Szervezetének (UNESCO) Bioszféra Rezervátuma (1978/’79, AT majd HU), és 2001 óta világörökség része (Dinka et al., 2016). Vizének utánpótlása túlnyomórészt csapadékból (1967–2012 közötti átlag: 181 × 10^6 m^3 év^{-1}; (Kubu et al., 2014)) és két kisebb vízfolyásból, a Wulka- (14%) és a Rákos-patakból történik. A tó partvonalának 10%-át érinti intenzív közvetlen antropogén tevékenység. Szinte folyamatos légmozgás jellemzi a területet, és az orográfia miatt északnyugati–északi a domináns irány (Józsa et al., 2008). A tó 80-100 évente kiszáradt az évszázadok folyamán (Löffler, 1979; Pannonhalmi, 2016), ennek elkerülésére vízkészlet-gazdálkodási intézkedéseket (pl. lefolyás-szabályozást) kellett bevezetni, ami nagyban befolyásolta a tavi nádasállományt (Dinka és Szeglet, 1998). Ez jelenleg a tó felületének ~56%-át teszi ki (Magyar et al., 2013) (**4.2-1C ábra**).

Mivel a tó frekventált rekreációs célpont, a mikrobiológia-higiénés állapota kulcskérdés, amelyet elsősorban négy szennyvíztisztítóból a nádason át közvetlenül a tóba jutó tisztított szennyvíz határoz meg (Kirschner et al., 2014). További három szennyvíztisztító a Wulka-patak vizén keresztül juttatja a Fertőbe a tisztított szennyvizet (Magyar et al., 2013). Ezen befolyók közül a podersdorfi a legjelentősebb (**4.2-1C ábra**), mely kifolyója egy kb. 300 m-es nádassáv előtt érkezik a tóba. Nagyobb vihareseményekkor a szennyvíztisztító csatornája túlfolyik, és bejut a tisztítatlan szennyvíz a tóba (Hatvani et al., 2018b; Herzig et al., 2019; Sommer et al., 2018). Ezen felül kisebb csatornákból is érkezhet állattartásból eredő fekáliás szennyvíz stb.

A fekáliás szennyezések feltárása az EU Bathing Water Directive évi öt mintázást ír elő a nyári rekreációs szezonban (EC, 2006a). Ennél azonban időben sokkal sűrűbb mikrobiológiai mintavételezés szükséges, hogy meghatározhatók legyenek a felételezett fekáliás forrópontok (Herzig et al., 2019) a tavon, és esetleges további gócokat is meg lehessen határozni. A Fertő Osztrák Vízügyi Mérnökségének azonban rendelkezésére áll 26 mintavételi pontról egy több évtizedes idősor a sztenderd fekál-indikátorbaktériumokról (**3.1-1 táblázat**), amely kombinálva általános vízminőségi és egyéb környezeti változókkal, megfelelő alapot adhat a Fertő fekáliás forrópontjainak feltárásához és egy új mintavételezési stratégia létrehozásához.

A Fertő fekáliás forrópontjainak vizsgálatához kapcsolódó kutatási célok

A Fertő esetében a kutatás célja volt, hogy sokváltozós adatelemző és geostatisztikai módszerekkel (i) meghatározzam a fekáliás forrópontokat a Fertőn mért SFIB-csíraszámértékek alapján, (ii) elkülönítsem a fekáliás szennyeződés fő hajtótényezőit, és (iii) hozzájáruljak a tó térbeli mintavételezési gyakoriságának optimalizálásához. Mivel ez a megközelítés egyedi a sekély tavak fekáliás szennyeződéseinek vizsgálatában, így bízom benne, hogy a kutatás példaként szolgálhat hasonló esetekben.

4.2.2. A Fertő fekáliás szennyeződésének vizsgálatából kapott eredmények

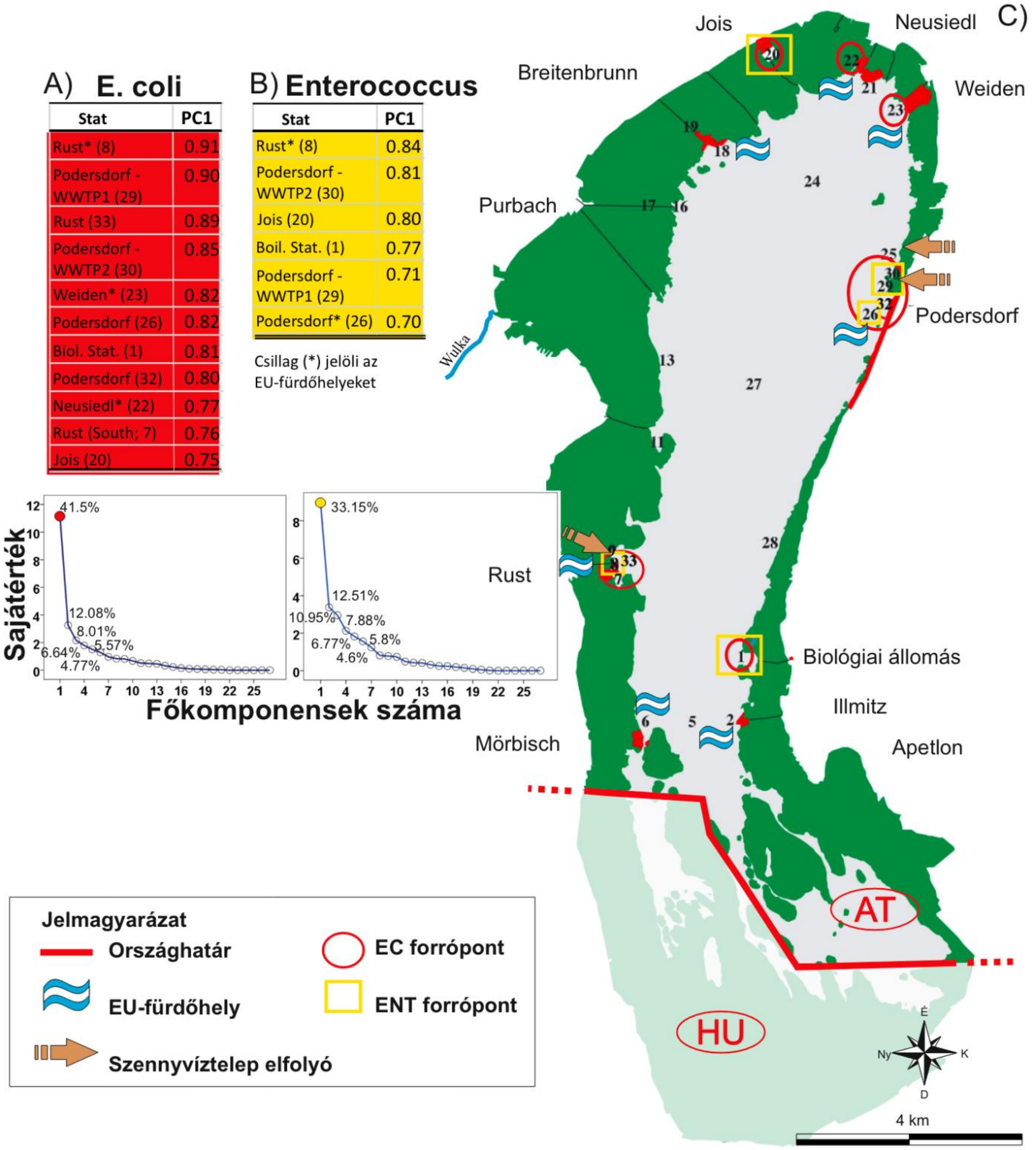
Osztrák kollégák felkérésére többváltozós adatelemzéssel modelleztem a Fertő osztrák részének fekáliás és vízgeokémiai adatait, hogy meghatározzuk, hol vannak a tavon a fekáliás forrópontok, és milyen antropogén tevékenység állhat ezek hátterében, hiszen ahogy ezt a 4.2.1-es fejezetben is írtam, e probléma részletes vizsgálata halaszthatatlanná vált a tó szempontjából. Itt jegyzem meg, hogy a Balatonon több mint 140 hivatalos EU-fürdőhely van kijelölve (EEA, 2023), amelyeknek, akárcsak a Fertőn, fekáliás szennyeződésre vonatkozóan rendszeresen ellenőrzik a vízminőségét. A téma fontossága ellenére mindösszesen csak egy-két tanulmány született a témában pl. Bíró et al. (1981). A Fertőn a jelen disszertációban bemutatott tanulmány az első a témában, pedig azokon a felszíni víztesteken, amelyeket rekreációs célzattal rendszeresen használnak, nem elegendő pusztán megállapítani esetleges vízminőségi anomáliákat, és a „fióknak dolgozni", hanem azokat – a vízi és a víztestet körülvevő környezet állapotát figyelembe véve – részletekbe menően ki is kell vizsgálni, és intézkedni.

A kutatás során a függő „response" változó az egy-egy mintavételi ponton mért SFIB-csíraszám volt. A PCA-hoz létrehoztam egy-egy korrelációs mátrixot az E. coli- és az Enterococcus-csíraszámok különböző mintavételi pontokon mért éves átlagaiból. Az így kapott főkomponenssúlyok a mintavételi pontok súlyát adták meg a kapott főkomponenseken belül, mindkét változóra egyenként (**4.2-1A-B ábrák**). Akárcsak a korábbi esetekben, a főkomponensek figyelembevételénél azt vizsgáltam, hogy azok sajátértéke egy felett legyen (Kaiser, 1960), és hogy a „könyökdiagramjuk" milyen lefutású (Cattell, 1966). Az így meghatározónak választott első komponensek a tó teljes eredeti E. coli- és Enterococcus-varianciájának rendre 41,5%-át (**4.2-1A ábra**) és 33,1%-át (**4.2-1B ábra**) magyarázták. Ezen főkomponensekben azokat a mintavételi pontokat, amelyek súlya nagyobb, mint 0,7, fekáliás forrópontnak neveztem, mivel ezekre összpontosult a korábban meghatározott fekáliás variancia tekintélyes része. *Ezzel meghatároztam a Fertő fekáliás forrópontjai* (**4.2-1C ábra**).

Következő lépésként megvizsgáltam azon külső környezeti magyarázó tényezőket, amelyek leginkább befolyásolhatták a forrópontokon az SFIB értékeit (Hatvani et al. (2018b): 'Relationship between the EC and ENT hotspots and the independent variables' fejezet). A szignifikáns (α = ,1) magyarázó változók a víz- és levegő-hőmérséklet, a napos órák száma, a Ruszton mért csapadék, a NO_3-N és az északi szélirány voltak (Hatvani et al. (2018b): 2. és 3. táblázat).

Félvariogram-vizsgálattal végül megállapítottam a SFIB-paraméterek térbeli autokorrelációját, hogy e két paraméterre vonatkozóan jobb képet kapjunk a tó lefedettségéről a térbeli mintavételezési gyakoriság szempontjából. Az eredmények alapján a mintavételi pontoknak kb. 1 km-es hatástávolsága van, tehát egy esetleges interpoláció bizonytalan eredményeket adna, hiszen jól látszik a térképről, hogy a meghatározott forrópontok elsősorban közvetlen környezetükről szolgáltatnak megbízható térbeli információt (Hatvani et al. (2018b): 4. ábra).

A forrópontokon megjelenő megemelkedett SFIB-csíraszámmal jellemezhető eseményeket és azok kapcsolatát a tavon észlelt viharesményekhez részletesen értelmeztük – elsősorban empirikusan – Herzig et al. (2019)-ben. Ennek a német nyelvű tanulmánynak kifejezetten az volt a célja, hogy közérthető formában hozzuk a helyi döntéshozók tudtára, hogy a probléma mennyire jelentős.



4.2-1. ábra. A Fertő, annak vizsgált mintavételi pontjai (számmal jelölve) és a meghatározott fekáliás forrópontok. Az első főkomponensek ’loading’-táblázata (0,7-es loading-érték fölött) és a főkomponens-analízis alatt kapott könyökdiagramok E. coli- A) és Enterococcus- B) csíraszámokra vonatkozóan. A térképen az E. coli-forrópontok piros körrel, míg az Enterococcus-forrópontok sárga négyzettel voltak jelölve C). A könyökdiagramokon a pontok feletti számok a magyarázott variancia mértékét adják meg az 1-nél nagyobb sajátértékű főkomponensekre. A zöld területek a Fertő partvonala mentén a nádasállományt jelzik. További részletek a 4.2.1-es fejezetben és a vonatkozó tanulmányokban (Hatvani et al., 2018b; Herzig et al., 2019; Magyar et al., 2013) olvashatók. Ábra Hatvani et al. (2018b)-ből került lefordításra.

Összefoglalva, a főkomponens-analízis egy új megközelítésű alkalmazásával meghatároztam a Fertő fekáliás forrópontjait és ezek hajtótényezőit, amelyek elsősorban településekhez, ill. az ottani antropogén tevékenységhez köthetők. Ilyen forrópontok pl. Rust (Ruszt) és a tavon álló hétvégi házak, amelyek környékén csapadékeseménykor csökken a szennyezés mértéke, feltehetőleg azért, mert a tulajdonosok akkor nem mennek ki a nyaralójukba, és nem ürítik illegálisan mellékhelységeiket a tóba; Podersdorf (Pátfalu), ahol vihareseménykor az északi szél rátolja a vizet a szennyvíztisztító kifolyójára, túltölti azt, és így a szennyvíz bekerül a tóba (Hatvani et al., 2018b; Herzig et al., 2019), továbbá SFIB a felkevert üledékből is visszajut a vízbe; vagy Weiden (Védeny) és Jois (Nyulas), ahol a gyakorta megemelkedett SFIB-érték szintén a nyaralók nem megfelelő szennyvízelvezetésére vezethető vissza. Ezzel szemben a nyílt vízen nem találtunk SFIB-forrópontot (Hatvani et al., 2018b), és csak alkalmanként jelent meg fekáliás szennyezés (Herzig et al., 2019). G*eostatisztikai vizsgálatokkal pedig megállapítottam, hogy a jelenlegi mintavételi hálózattal nem ajánlott interpolálni SFIB-re,* így javaslatom, hogy egy intenzív térbeli mintavételezési kampánnyal reprezentatívan mérjék fel a tó különböző élőhelyeit és az antropogén tevékenységnek fokozottan kitett területeit, figyelembe véve a tó áramlási viszonyait és az uralkodó északi szélirányt (Antal et al., 1991; Józsa et al., 2008). Az így kapott adatok alapján a tavon egy térben hatékonyabb és SFIB-re reprezentatívabb mintavételezési hálózat tudna üzemelni (Sommer et al. (2018)). Az eredmények felhívták a figyelmet több, feltételezhetően jogsértő és környezetkárosító tevékenységre, amelyek megszüntetésével az UNESCO világörökségi védelme alatt is álló Fertő vízminősége és a helyi turizmus nagyban fog profitálni.

4.3. A modern csapadék δ^2H- és δ^{18}O-értékei közötti lineáris kapcsolat alakulása a Mediterráneumban

Míg az előző fejezetekben bemutatott monitoringhálózatokból származó adatok egy-egy térben jól lehatárolható víztest vízminőségi állapotára adtak betekintést, az alábbi fejezetekben tárgyalt környezeti adatok szubkontinentális léptékű mintavételi hálózatokból származnak: a Mediterráneumból (4.3. fejezet) és nyugat-antarktiszi jégfuratokból (4.4. fejezet). E területekről származó vízminták (csapadék és jégfurat) stabilizotóp-összetételének térbeli, geomatematikai modellezésével fogok egy-egy tanulmány/fejezet erejéig részletesebben foglalkozni.

4.3.1. Lokális csapadékvízvonalak (LCsVV-k) alkalmazásának bemutatása a Mediterráneumban

A mediterrán térségben egyre nagyobb léptéket ölt a vízhiány (García-Ruiz et al., 2011; Hoerling et al., 2012; Iglesias et al., 2007). Az izotóphidrológia eszközeiben pedig jelentős a potenciál a regionális vízkörforgást irányító folyamatok jobb megértésének segítésében, így a tudásalapú vízkészlet-gazdálkodás előmozdításában (Aggarwal et al., 2005). Az LCsVV a δ_p jellegzetes eloszlását írja le, és amennyiben hosszú távú megfigyelésekből határozzák meg, referenciaként szolgálhat izotóphidrológiai alkalmazásokban, mint pl. a beszivárgási időszak becslése, vagy a talajvízből kinyert izotópos jel paleoklimatológiai értelmezése (Clark és Fritz, 1997). A mediterrán térségben is számos izotóphidrológiai tanulmány alkalmazta az LCsVV-ket felszín alatti (Brkić et al., 2020; Christofi et al., 2020; Hssaisoune et al., 2022; Koeniger et al., 2017; Liotta et al., 2013; Trabelsi et al., 2020; Túri et al., 2020; Vasić et al., 2019), vagy felszíni hidrológiai problémák kezelésénél (El Ouali et al., 2022; Gat és Dansgaard, 1972; Serianz et al., 2021). Azonban a viszonylag távoli mintavételi pontok adataiból meghatározott LCsVV-k száma az utóbbi időben megnőtt (pl. Boumaiza et al. (2021); ElGhawi et al. (2021); Túri et al. (2019)), jelezvén, hogy a régió izotópos hidrológiai kutatásaiban nagy az igény ilyen referenciaadatokra. Az LCsVV-k megbízható becslése így azon területeken is lényeges, ahol nem folyik izotópmeteorológiai monitoring. Továbbá meg kell jegyezni, hogy egy adott vizsgálat helyéhez képest távoli ponton mért adatokból meghatározott LCsVV-k félrevezető információt adhatnak helyi vízgazdálkodási alkalmazásokban, pl. Serianz et al. (2021). A mediterrán térségben, regionális kampányoknak (IAEA, 2005) és további szubregionális megfigyeléseknek hála, a térbeli adatlefedettség nagymértékben megnőtt (Hatvani et al. (2023): 1B. ábra), azonban ez még mindig nem elegendő számos tudományos kérdés megválaszolásához.

Egy nemrégiben készült tanulmány foglalkozott az LCsVV-k meredekségének térbeli alakulásával Európa mérsékelt égövi és boreális régióiban (Lécuyer et al., 2020). Azonban nem foglalkozott az LCsVV-k meredekségének meghatározásával a mediterrán térségre vonatkozóan, ahogyan az LCsVV-k tengelymetszetének modellezésével sem.

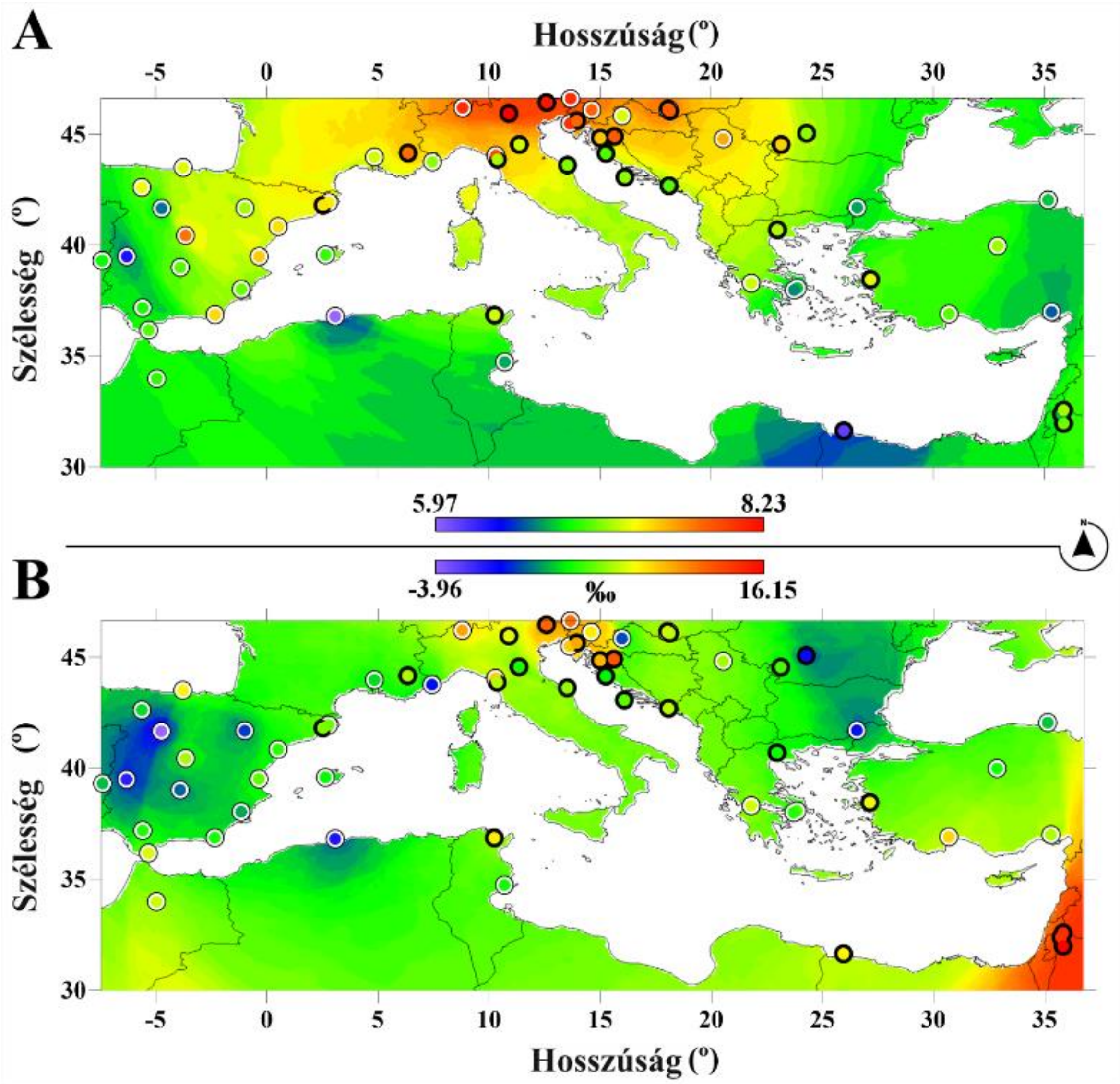
A modern mediterrán csapadék lokális vízvonalainak térben folytonos modellezésének célja

A munkám célja volt, hogy (i) kiegészítse a Lécuyer et al. (2020) által vizsgált területet a mediterrán térséggel, nem csak meredekség, hanem (ii) az LCsVV-k tengelymetszetének tekintetében is. Emellett a Hatvani et al. (2023)-ban bemutatott eredmények módszertani kiterjesztést is jelentenek, mivel (iii) a Mediterrán térség LCsVV-k meredekségének és tengelymetszetének értékét klasszikus statisztikai és gépi tanulási módszerekkel is becsültem, és ezek teljesítményét összesen 14 különböző megközelítésben (Hatvani et al. (2023): 2.3.3. fejezet) hasonlítottam össze, figyelembe véve a régió éghajlati és földrajzi adottságait.

4.3.2. A modern mediterrán LCsVV-k térben folytonos becsléséből kapott eredmények

Kollégákkal, doktorandusz és szakdolgozó hallgatómmal 249 mediterrán csapadékmegfigyelő állomás hosszú távú δ_P-adatait vizsgáltam meg. Ezek közül a kiválasztott fókuszidőszakban (2000–2015) 62 felelt meg az alkalmazott szigorú adatszűrési feltételeknek; Hatvani et al. (2023): 2.3.1. fejezet. Az RMA-regresszióval meghatározott LCsVV-k meredekségét és tengelymetszetét a vizsgált 14 eljárás közül (lásd 3.2.1. és Hatvani et al. (2023): 2.3.3. fejezete) az állomások távoságmátrixának és tengerszint feletti magasságának figyelembevételével végzett RF becsülte meg a legpontosabban; Hatvani et al. (2023): 3.2. fejezet és 3. ábra.

Az LCsVV-k meredeksége és tengelymetszete rendre ~5,9 és 8,2, illetve -3,9 és 16,1‰ között változott. A tengelymetszet 7 alatti értékeket is felvett a Mediterráneum nyugati peremén, valamint Észak-Afrikában. A legmagasabb, 8 feletti értékek északon, az olasz Alpok lábainál voltak (**4.3-1A ábra**). Elmondható, hogy egy folytonos változás helyett szubregionális mintázatokat mutat a mediterrán LCsVV-k interpolált térképe. A vizsgált terület nyugati peremétől egészen az Adriai-tengerig (~12°E) hasonló tendencia figyelhető meg, mint amit Európa mérsékelt égövi és boreális régióira modelleztek, elérve a 8-as meredekséget; Lécuyer et al. (2020): 4. ábra. Kelet felé továbbhaladva azonban ez megváltozik, és míg Lécuyer et al. (2020) folyamatos 9-es meredekségig tartó növekedést jelentett, a vizsgált régió esetében csökkenést tapasztaltam (**4.3-1A ábra)**.



4.3-1. ábra. A vizsgált mediterrán makrorégióba modellezett LCsVV-k meredekségének (A) és tengelymetszetének (B) pontbeli értékei és a RF$_{sp\text{-}ele}$ modellel kapott interpolált térképek. A színes pontok fehér körvonallal az 1. kategóriába tartozó csapadékmérő állomásokat, míg a fekete körvonal a 2. kategóriába tartozó állomásokat jelölik. További részletek a 4.3.1. fejezetben és Hatvani et al. (2023)-ban olvashatók. Ábra Hatvani et al. (2023) szerint.

A tengelymetszet térbeli alakulása még összetettebb mintázatot mutat. Ibéria középső részén < -3,5‰, míg a part menti részeken 2,7 és 6‰ közötti mozog. Egy fokozatos emelkedés után az Adriai-tenger környékén már magasabb értékeket mutat, amelyek Északkelet-Olaszországban és Szlovéniában a 10‰-et is meghaladják (**4.3-1B ábra**). Ez egybeesik a legnagyobb meredekségek helyével (**4.3-1A ábra**). Az Adriai-tengertől keletre fokozatos csökkenés figyelhető meg: a keleti hosszúság 17°-a körüli 7,5‰-es tengelymetszetek 5‰ körülire csökkennek (keleti hosszúság 20–23°). A Fekete-tenger térségében tapasztalható szubregionális mintázat pedig < 3,2‰ tengelymetszetet mutat. Elmondható, hogy a kisebb eltérések ellenére is a bemutatott modell összességében jól megegyezik korábbi helyi/regionális becslésekkel; Hatvani et al. (2023): 3.3. fejezet.

A kiszámított modellek eredményeinek összevetése és független ('out-of-sample') validációja alapján *bemutattam, hogy a mediterrán térség modern δ_p-adataiból RMA-*

regresszióval számolt LCsVV-k meredekségének és tengelymetszetének térbeli becslésére az állomások távolságmátrixának és tengerszint feletti magasságának figyelembevételével végzett RF eljárás a legalkalmasabb a figyelembe vett modellek közül.

Az LCsVV-k paramétereinek térben folytonos becslésével bemutattam, hogy (i) a mediterrán térségben számos szubregionális és regionális mintázat jelenik meg, mind meredekség, mind tengelymetszet tekintetében. Ezek pedig rávilágítanak annak fontosságára, hogy a csapadékvízvonalakra vonatkozóan finom skálájú, térben folyamatos becsléseket kell levezetni, és nem pedig számos, egymástól több száz vagy akár több ezer kilométerre lévő állomás adatait kell összesíteni. (ii) Mivel a Nyugat-Mediterráneumban nem volt homogén a mintázata az LCsVV-k paramétereinek, az eredmények kétségbe vonják egy egységes regionális csapadékvízvonal meghatározását. Ez pedig megkérdőjelezi a korábban definiált nyugat-mediterrán csapadékvízvonal „WMMWL" mint izotóphidrológiai referencia (Celle-Jeanton et al., 2001) további használhatóságát a modern csapadékra.

A mediterrán csapadék δ^2H – δ^{18}O statisztikus kapcsolatának bemutatott makroregionális modellje térben folyamatos becslést ad a modern (2000 utáni) helyi csapadékvízvonalak meredekségére és tengelymetszetére, frissítve a huszadik századból származó szubregionális tanulmányok eredményeit. Javasolt tehát a közeli, hosszú távú megfigyeléssel rendelkező állomásokból származó helyi becslések, vagy közeli állomás hiányában a jelenlegi interpolált termékből nyert becsléseket használni a régióra, amelyek egy térinformatikai adathalmaz formájában, az eredeti tanulmány függelékeként elérhetőek itt.

4.4. δ^2H- és δ^{18}O-értékek alakulása felszíni hó- és firnrétegekben egy antarktiszi makrorégióban

4.4.1. Csapadék hidrogén- és oxigén-stabilizotópos összetétele az Antarktiszon – háttér és globális jelentőség

A földi hidroklíma évtizedes–évszázados természetes változékonyságának megértéséhez elengedhetetlen, hogy jobb képet kapjunk a múltbéli – főként az ipari forradalom előtti – időkről. A jégfuratok évszakostól (Hammer, 1989; Kuramoto et al., 2011) százezer éves léptékekig szolgáltatnak információt a Föld klímájáról (Jouzel, 2013). A megfigyelt csapadék δ^2H- és δ^{18}O-értékek és globális éghajlati modellek eredményeinek integrált kiértékelésével pontosabb képet kaphatunk a hidroklimatológiai változásokat hajtó földi klimatikus tényezőkről (Werner és Heimann, 2002). Az antarktiszi jégfuratokból származó stabilizotóp-adatok kezelése jelentős kihívással járó feladat, mivel a magok térbeli elérhetősége ritka és nagyon változó a kontinensen, továbbá az időszak, amelyet lefednek, nagyban változhat (IPICS, 2006; Masson-Delmotte et al., 2008; Steig et al., 2005).

Annak ellenére, hogy az elmúlt évtizedben nagyfokú adatszintézisre került sor a paleoklíma-közösségen belül, a hidroklíma-mintázatok szintézisében csak korlátozott előrelépés történt (Konecky et al., 2020). Noha már rendelkezésre áll egy globális adatbázis az elmúlt 2000 év δ^2H- és δ^{18}O-rekordjairól, de ezek térbeli lefedettsége (Konecky és McKay, 2020) továbbra is megköveteli a pontok közötti interpolációt. A csapadékvíz izotóp-összetételére számos globális (Terzer-Wassmuth et al., 2021; Terzer et al., 2013; van der Veer et al., 2009) tájkép létezik, de ezek egyike se foglalta magába az Antarktiszt. Egyedül a felszíni hórétegeket vizsgálva hoztak létre egy hasonló térképet, de az nem fedi a jelen dolgozatban is vizsgált antarktiszi selfterületeket (Wang et al., 2010).

Az izotópos tájképek térben folytonos becslést adnak a vizsgált izotópos változó értékeiről, ami pl. az ún. 'isotope enabled' globális éghajlati modellek teljesítményének értékelésénél szolgáltathat referenciaadatokat (Lachniet et al., 2016; Sturm et al., 2005), de alkalmazhatók olyan regionális légkörzési modellek ellenőrzésére is, amelyeknek fókuszában a hidrológiai ciklus és annak izotópokkal jellemezhető egységei vannak.

A nyugat-antarktiszi firn/jég oxigénizotópos tájkép létrehozásának célja

A nyugat-antarktiszi jégfurathálózat vizsgálatával az volt a célom, hogy (i) meghatározzam azon földrajzi tényezőket, amelyek elsődlegesen irányítják a régió jégfuratainak vízizotóp-összetételét, (ii) megvizsgáljam a jégfuratok vízizotóp-összetételének térbeli autokorrelációját, ami elengedhetetlen a geostatisztikai modellezéshez (Herzfeld, 2004), és (iii) létrehozzak egy regionális izotópos tájképet (Bowen, 2010), amely reprezentatívan írja le a régió vízizotóp-összetételének térbeli eloszlását.

4.4.2. (Nyugat-)antarktiszi jégfuratok stabilizotóp-összetétele geostatisztikai vizsgálatának eredményei

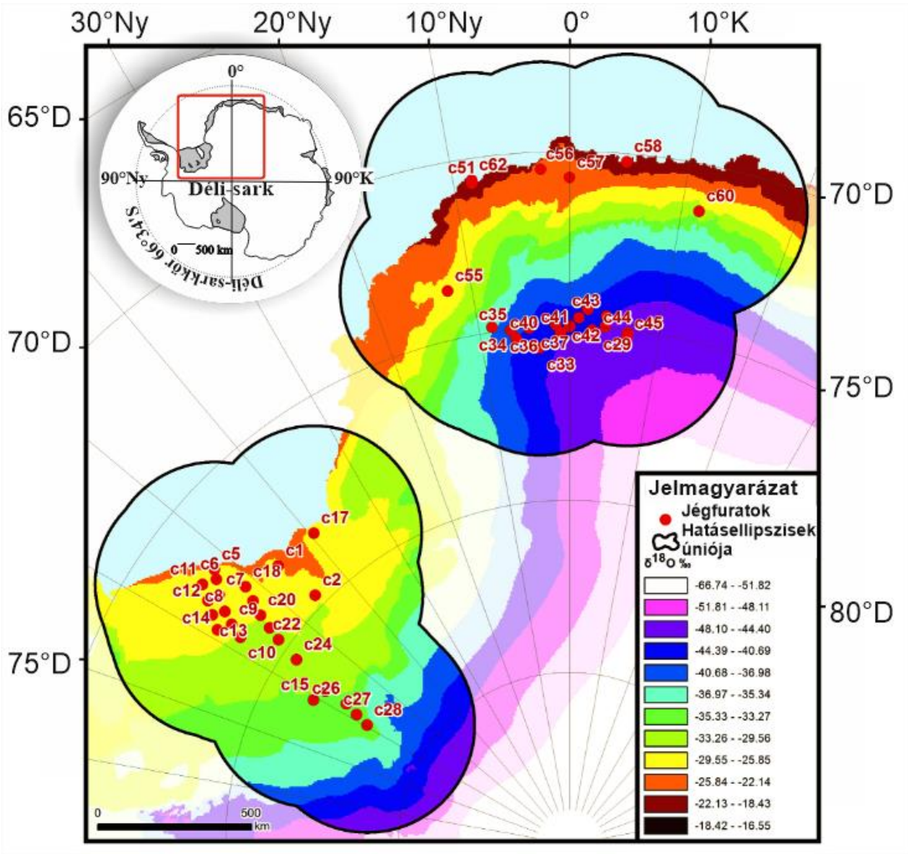
A (nyugat-)antarktiszi vizsgált terület jégfuratainak oxigénizotóp-összetételeire izotópos tájképet készítettem. Első lépésben többváltozós regresszióval meghatároztam a stabilizotóp-összetételt regionálisan meghatározó lehetséges földrajzi tényezők közül (Lorius és Merlivat, 1977; Masson-Delmotte et al., 2008) azokat, melyeknek a szerepe a térségben a legjelentősebb, hogy az azokból eredő trendet el tudjam távolítani az adatokból, így egy stacionárius adathalmazt kapjak, amely elengedhetetlen geostatisztikai vizsgálatokhoz (Füst és Geiger, 2010; Hohn, 1999). Ezt a furatok földrajzi hosszúság szerinti pozíciója (LON), tengerszint feletti magassága (ELE) és tengertől való távolsága (D) segítségével értem el ($R^{\sim 2}$ = 98%; **1. E.**), számtalan kombinációt megvizsgálva (Hatvani et al. (2017b): 3.1. fejezet), O'Brien (2007) ajánlása alapján az $R^{\sim 2}$, MAE, VIF és szignifikanciaszinteket is figyelembe véve.

$$\delta^{18}\hat{O} = -(20{,}51 \pm 0{,}57) + (2{,}64\ [\pm 0{,}57] \times 10^{-6}) \times LON - (5\ [\pm 0{,}27] \times 10^{-3}) \times ELE - (1{,}9\ [\pm 0{,}12] \times 10^{-5}) \times D \qquad \textbf{(1. Egyenlet)}$$

Az 1. egyenlettel (**1. Egyenlet**) kiszámítottam az ún. „initial gridet" és a mért értékektől vett eltérések képezték a maradéktagok ponthalmazát (Hatvani et al. (2017b): 4a ábra) amelyen a félvariogram-vizsgálatot és a krigelést (Hatvani et al. (2017b): 3.2. fejezet) végeztem el. Ennek eredményeképp kaptam az ún. „residual gridet", és ezek összegzésével hoztam létre végül az izotópos tájképet (Hatvani et al. (2017b): 3.3. fejezet).

A kapott regressziós együtthatók megadták a magassági hatás léptékét (0,005‰ m^{-1}), amely nagyságrendileg megegyezik az irodalomban találhatókkal (Hatvani et al. (2017b): 4.1. fejezet). A további két változó lényegi szerepe is egyaránt logikus, hiszen a tengertől való távolság befolyásolja a folyamatos csapadékkihullás mértékét, ahogy a légtömegek a kontinens belseje felé vándorolnak. Ez azonban nem magyarázza a „másodrendű" irányító hatásokat, mint pl. a specifikus, lokális légtömeg-trajektóriákat, vagy az évszakosságból adódó eltéréseket (Sodemann és Stohl, 2009). További részletekért lásd Hatvani et al. (2017b): 4.1. fejezet.

A geostatisztikai modellezés során a félvariogram-vizsgálatot „autofittinggel" végeztem el, amely egy gaussi elméleti félvariogram illesztését javasolta (C_0 = 0,13; $C_0 + C$ = 1,27; a = 350 km; r^2 = 0,72; bin-méret ~55 km; Hatvani et al. (2017b): 4b ábra, paraméterek magyarázata annak ábraaláírásában található). A függvényt a „residual grid" krigelésénél súlyfüggvényként használtam. Az intepolált „residual gridet" összeadtam az „intial griddel", így kaptam meg a Nyugat-Antarktisz interpolált izotópos δ^{18}O-tájképét (**4.4-1. ábra**) egy térinformatikai adathalmaz formájában, amely elérhető az eredeti tanulmány függelékeként itt. A tájképen jelöltem a furatok térbeli hatástávolságának unióját is, amely kijelöli a jelenlegi hálózat által le nem fedett területeket. A „residual grid” krigelési szórása (’kriging standard deviation’) > 0,45 ‰ (Hatvani et al. (2017b): 4c,d ábra), amit a furatok térbeli elhelyezkedése okoz (Chilès és Delfiner, 2012), de nem is várhatnánk a természetben, hogy ez nulla legyen, mint ideális esetben (Wackernagel, 2003). Mindemellett e bizonytalanságok elfogadhatóak, mivel a δ^{18}O-analízis szokásos analitikai pontossága izotópanalitikai módszertől függően ~0,1 - 0,2‰ (Lis et al., 2008).



4.4-1. ábra. A vizsgált antarktiszi makrorégió (bal felső térkép piros négyzet) és a geostatisztikai modellezéssel kapott δ^{18}O-izotópos tájkép (1970-1988). A 350 km-es hatásellipszisek unióját a fekete vonal jelöli. A piros pontok jelölik a felhasznált jégfuratok helyét. Átdolgozva Hatvani et al. (2017b) alapján.

A nyugat-antarktiszi makrorégióban található hó/firn oxigénizotóp-összetétel adatsorainak *vizsgálatával (i) kimutattam, hogy a vizsgált terület hó/firn* δ^{18}O értékek térbeli *varianciáját leginkább a földrajzi hosszúság, a tengerszint feletti magasság és a parttól való távolság határozza meg. (ii) Lehatároltam azon területeket a régión belül, amelyeket a jelenlegi jégfurathálózat hó/firn* δ^{18}*O összetételének térbeli varianciája (a 350 km-es térbeli hatástávolsággal) nem fed le.* Ezek azok a területek, amelyek az optimális helyszínválasztás ismert kritériumain felül (Vance et al., 2016) geostatisztikai szempontból leginkább hozzájárulhatnak a jelenlegi jégmaghálózat lefedettségéhez, amennyiben itt kerülnek új jégmagok mélyítésre. Ezen felül az eredmények minden eddiginél teljesebb képet nyújtanak a terület hó/firn oxigénizotóp-összetételének varianciájáról, hiszen a vizsgált régió 20%-a selfterületeken van, amelyek mindeddig kiestek a kontinentális δ^{18}O-térképekből (Wang et al., 2009; Wang et al., 2010), és globális δ^{18}O-tájképekből (Terzer et al., 2013; van der Veer et al., 2009). A kapott jég/firn δ^{18}O-tájképe kiegészítő bemeneti adat lehet az „isotope enabled” globális cirkulációs modellek esetében, amikor nincs elegendő rendelkezésre álló mért csapadék δ^{18}O-adat.

5. ÖSSZEFOGLALÁS

A Zala–Kis-Balaton Vízvédelmi Rendszer (KBVR)–Balaton, mint kis vízhozamú vízfolyás–eutróf tó–wetland–sekély tó kaszkádrendszer fő bemeneti pontján a Zalán a diffúz mezőgazdasági terheléshez és jellemzően pontforrás eredetű szennyvízhez köthető folyamatok szerepét átvették a természeti folyamatokat inkább jellemző háttértényezők (pl. vízhőmérséklet). A tapasztalt jelenség hátterében elsősorban a vízminőség-védelmi intézkedések és az antropogén tevékenység csökkenése, az 1990-es években bekövetkezett nagymértékű mezőgazdasági átrendeződés áll. Módszertani szemszögből a Zala vízgyűjtőjéről érkező terhelések és azok megjelenése a KBVR-en, beleértve a vízgyűjtőre ható külső tényezőket, időben eltolva jelentek meg, és az ilyen folyamatok vizsgálata a faktoranalízis egy speciális esetét, a dinamikus faktoranalízist követelte meg, mivel ez képes figyelembe venni az eltolt korrelációs struktúrákat, amelyek a vizsgálat tárgyát képezték.

A KBVR különböző élőhelyein az elsődleges tápanyagformák jelentősen eltérő mértékben mutattak hidro(geo)kémiai évszakosságot. Az eutróf tóban zajló folyamatok voltak képesek a legjelentősebb mértékben tükrözni a meteorológiai változók évszakosságát, a makrofita dominálta wetland térrész (Ingói-berek) pedig szétvált egy ún. „zavartalan” és „zavart wetlandre”. Utóbbi a Zalához hasonló viselkedést mutatott a hidro(geo)kémiai évszakosság szempontjából a rendszertelenül beérkező 40%-nyi vízhozamtöbblet miatt, amelyet a Marót-völgyi-csatorna, a Zala-Somogyi-csatorna és az Egyesített-övcsatorna hoznak a rendszerbe.

A Balaton három évtizednyi vízminőségét és egyben trofikus állapotát leíró adatainak új megközelítésű elemzése alapján az eddigi irodalommal egybecsengő kép rajzolódott ki. A tó trofikus állapota a nyugati medencék hipertróf állapotából a Siófoki-medencében már oligotróf állapotokat is mutatott klorofill-a alapján. Ez a javuló trend pedig az egész tóra jellemző az 1990-es évek közepétől kezdődően, amely azonban időben késleltetve jelent csak meg a medencékben kelet felé haladva. Ezt tükrözi három szignifikánsan elváló vízminőségi időszak (1985–1994, 1995–2003 és 2004–2017). A kelet felé csökkenő szerepe a tápanyagoknak és a Zalából érkező ÖP-terhelés folyamatosan gyengülő meghatározó jellege is jelzi, hogy a medencék térben is máshogyan reagáltak a terhelésekre.

A Fertő fekáliás forrópontjai és ezek hajtótényezői elsősorban településekhez, ill. az ottani antropogén tevékenységhez köthetők. Ilyen forrópontok pl. Rust és a tavon álló hétvégi házak, amelyek környékén csapadékeseménykor csökken a szennyezés mértéke; Podersdorf, ahol vihareseménykor az északi szél rátolja a vizet a szennyvíztisztító kifolyójára, túltölti azt, és így a szennyvíz bekerül a tóba, továbbá a sztenderd fekáliás indikátorbaktériumok (SFIB) a felkevert üledékből is visszajutnak a vízbe; vagy Weiden és Jois, ahol a gyakorta megemelkedett SFIB-érték szintén a nyaralók nem megfelelő szennyvízelvezetésére vezethető vissza. Ezzel szemben a nyílt vízen nem jelent meg SFIB-forrópont, és fekáliás szennyezés is csak alkalmanként. A geostatisztikai vizsgálatok kimutatták, hogy a jelenlegi mintavételi hálózattal nem ajánlott interpolálni SFIB-re. Javasolt egy intenzív térbeli mintavételezési kampánnyal reprezentatívan felmérni a tó különböző élőhelyeit és antropogén tevékenységnek fokozottan kitett területeit, figyelembe véve a tó áramlási viszonyait. Az így kapott adatok alapján a tavon egy térben hatékonyabb és SFIB-re reprezentatívabb mintavételezési hálózat tudna üzemelni. Az eredmények felhívták a figyelmet több, feltételezhetően jogsértő és környezetkárosító tevékenységre, amelyek megszüntetésével az UNESCO világörökségi védelme alatt is álló Fertő vízminősége és a helyi turizmus nagyban fog profitálni.

A mediterrán térség modern csapadék stabilizotópos adataiból ’reduced major axis’ regresszióval számolt lokális csapadékvízvonalak (LCsVV-k) meredekségének és tengelymetszetének térbeli becslése a mediterrán csapadék-mérőállomások távoságmátrixának és tengerszint feletti magasságának figyelembevételével végzett ’random forest’ eljárás a legalkalmasabb a figyelembe vett 14 modell közül. Miután megtörtént az LCsVV-k becslése bemutattam, hogy a Földközi-tenger tágabb medencéjében számos szubregionális és regionális mintázat jelenik meg, mind meredekség, mind tengelymetszet tekintetében. Ezek pedig rávilágítanak annak fontosságára, hogy a csapadékvízvonalakra vonatkozóan finom skálájú, térben folyamatos becsléseket kell levezetni, és nem pedig egymástól több száz vagy akár több ezer kilométerre lévő állomás adatait kell összesíteni. Mivel a Földközi-tenger nyugati medencéjében nem volt egységes az LCsVV-k paramétereinek mintázata, az eredmények kétségbe vonják egy regionális csapadékvízvonal meghatározását. Ez pedig megkérdőjelezi a korábban definiált nyugat-mediterrán csapadékvízvonal mint izotóp-hidrológiai referencia (Celle-Jeanton et al., 2001) további használhatóságát a modern csapadékra.

A nyugat-antarktiszi makrorégióban található hó/firn oxigénizotóp-összetétel-adatsorainak vizsgálatával kimutatta, hogy a terület δ^{18}O-varianciáját leginkább a földrajzi hosszúság, a tengerszint feletti magasság és a parttól való távolság határozza meg. A geostatisztikai modellezés felfedte azon területeket a vizsgált régión belül, amelyeket a jelenlegi jégfurathálózat térbeli varianciája (350 km-es hatástávolság) nem fed le.

KÖSZÖNETNYILVÁNÍTÁS

Elsősorban köszönettel tartozom Demény Attilának, Kovács Józsefnek, és Kern Zoltánnak, akikhez bármikor fordulhattam kérdéseimmel és akik mindig segítő kezet nyújtottak kutatásaimban.

Az értekezést elsősorban egyes szám első személyben írtam, és azon eredményekre fókuszáltam benne, amelyeket első, vagy alkalmanként levelező szerzőként jegyzett tanulmányokban publikáltam, azonban *ezen munkák mindegyike csapatmunkában, kollégákkal valósult meg, amiért szeretnék nekik köszönetet mondani*, hiszen az eredmények a közös erőfeszítések nélkül nem jöhettek volna létre. Név szerint: Ala Eddine Smati, Alexander Farnleitner, Alexander Kirschner, Alois Herzig, Clement Adrienne, Hoffmann Richárd, Kern Zoltán, Kohán Balázs, Korponai János, Kovács József, Kovácsné Székely Ilona, Marcus Leuenberger, Márkus László, Polona Vreča Szatmári Gábor, Tanos Péter, és Vinicius de Barros. Ki szeretném közülük emelni Kovács Józsefet, Kern Zoltánt és Clement Adriennet, akikkel e dolgozat gerincét adó tanulmányokon a legtöbbet dolgoztunk, és akik szakmai hozzáértése, türelme és embersége nélkül ez a disszertáció nem készülhetett volna el.

Köszönettel tartozom hallgatóimnak, Ala Eddine Smatinak, Erdélyi Dánielnek, Topál Dánielnek és Vinicius de Barrosnak az együttes kutatómunkáért, az inspiráló légkörért és adott esetekben segítségükért; hálás vagyok továbbá azért a szakmai segítségért, amelyet Dinka Máriától és Berczik Árpádtól, illetve Tanos Márton bölcsésztől kaptam, aki a dolgozatot nyelvileg lektorálta, és Paul Thatchertől, aki a dolgozat gerincét adó angol nyelvű tanulmányok lektorálását vállalta. Rendkívül hálás vagyok a szakmai munkám során elnyert kutatási ösztöndíjakért és támogatásokért az Osztrák-Magyar Akció Alapítványnak, az MTA Bolyai János Kutatási Ösztöndíj Kuratóriumának, az Emberi Erőforrások Minisztériumának, valamint a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatalnak. Végül, de nem utolsósorban szeretném köszönetem kifejezni családomnak a türelemért, amit egy-egy intenzív kutatással töltött nap után tanúsítottak.

IRODALOMJEGYZÉK

A rövid értekezés alapjául szolgáló közlemények (PhD fokozatszerzés óta)

Hatvani, I.G., Kovács, J., Márkus, L., Clement, A., Hoffmann, R., Korponai, J., 2015. Assessing the relationship of background factors governing the water quality of an agricultural watershed with changes in catchment property (W-Hungary). Journal of Hydrology 521: 460-469. https://doi.org/10.1016/j.jhydrol.2014.11.078

Hatvani, I.G., de Barros, V.D., Tanos, P., Kovács, J., Székely Kovács, I., Clement, A., 2020a. Spatiotemporal changes and drivers of trophic status over three decades in the largest shallow lake in Central Europe, Lake Balaton. Ecological Engineering 151: 105861. https://doi.org/10.1016/j.ecoleng.2020.105861

Hatvani, I.G., Smati, A.E., Erdélyi, D., Szatmári, G., Vreča, P., Kern, Z., 2023. Modeling the spatial distribution of the meteoric water line of modern precipitation across the broader Mediterranean region. Journal of Hydrology 617: 128925. https://doi.org/10.1016/j.jhydrol.2022.128925

Hatvani, I.G., Clement, A., Korponai, J., Kern, Z., Kovács, J., 2017a. Periodic signals of climatic variables and water quality in a river – eutrophic pond – wetland cascade ecosystem tracked by wavelet coherence analysis. Ecological Indicators 83: 21-31. https://doi.org/10.1016/j.ecolind.2017.07.018

Hatvani, I.G., Kirschner, A.K.T., Farnleitner, A.H., Tanos, P., Herzig, A., 2018b. Hotspots and main drivers of fecal pollution in Neusiedler See, a large shallow lake in Central Europe. Environmental Science and Pollution Research 25(29): 28884-28898. https://doi.org/10.1007/s11356-018-2783-7

Hatvani, I.G., Leuenberger, M., Kohán, B., Kern, Z., 2017b. Geostatistical analysis and isoscape of ice core derived water stable isotope records in an Antarctic macro region. Polar Science, 13: 23-32. https://doi.org/10.1016/j.polar.2017.04.001

A szerző további hivatkozott, a rövid értekezés alapjául nem szolgáló közleményei

[5]*Csábrági, A. et al., 2019. Estimation of dissolved oxygen in riverine ecosystems: Comparison of differently optimized neural networks. Ecological Engineering, 138: 298-309.

Hatvani, I. et al., 2011a. Adatelemző módszerek alkalmazásának feltételei és lehetőségei a felszíni és felszín alatti víz védelmében. In: Bunyevácz, J., Csonka, P., Fodor, I., Gálosi-Kovács, B. (Eds.), A fenntartható fejlődés, valamint a környezet-és természetvédelem összefüggései a kárpát-medencében. MTA Pécsi Akadémiai Bizottság, Pécs.

Hatvani, I.G. et al., 2021. Geostatistical evaluation of the design of the precipitation stable isotope monitoring network for Slovenia and Hungary. Environment International, 146: 106263.

Hatvani, I.G., Clement, A., Kovács, J., Kovács, I.S., Korponai, J., 2014a. Assessing water-quality data: The relationship between the water quality amelioration of Lake Balaton and the construction of its mitigation wetland. Journal of Great Lakes Research, 40(1): 115-125.

[5] A *-gal megjelölt tanulmányokban utolsó és/vagy levelező szerző voltam.

Hatvani, I.G., Dokulil, M.T., Clement, A., 2022. The role of wetlands in mitigating impacts from diffuse agricultural loads. In: Mehner, T., Tockner, K. (Eds.), Encyclopedia of Inland Waters (second edition). Elsevier, Oxford, pp. 285-299. ISBN:9780128220412

Hatvani, I.G., Erdélyi, D., Vreča, P., Kern, Z., 2020b. Analysis of the spatial distribution of stable oxygen and hydrogen isotopes in precipitation across the Iberian Peninsula. Water, 12(2): 481.

Hatvani, I.G., Kern, Z., Leél-Őssy, S., Demény, A., 2018a. Speleothem stable isotope records for East-Central Europe: Resampling sedimentary proxy records to obtain evenly spaced time series with spectral guidance. Earth System Science Data, 10(1): 139-149.

Hatvani, I.G., Kovács, J., Kovács, I.Sz., Jakusch, P., Korponai, J., 2011b. Analysis of long-term water quality changes in the Kis-Balaton Water Protection System with time series-, cluster analysis and Wilks' lambda distribution. Ecological Engineering, 37(4): 629-635.

Hatvani, I.G., Magyar, N., Zessner, M., Kovács, J., Blaschke, A.P., 2014b. The water framework directive: Can more information be extracted from groundwater data? A case study of Seewinkel, Burgenland, Eastern Austria. Hydrogeology Journal, 22(4): 779-794.

Herzig, A. et al., 2019. Mikrobiologisch-hygienische Untersuchungen am Neusiedler See – von der Einzeluntersuchung zum Gesamtkonzept.. Österreichische Wasser- und Abfallwirtschaft, 71(11): 537-555.

*Kovács, J. et al., 2012b. Delimiting sub-areas in water bodies using multivariate data analysis on the example of Lake Balaton (W Hungary). Journal of Environmental Management, 110: 151-158.

*Kovács, J. et al., 2012c. Analysis of water quality data for scientists. In: Kostas Voudouris, Voutsa, D. (Eds.), Water quality monitoring and assessment. InTech, pp. 65-94. ISBN:978-953-51-0486-5

Kovács, J. et al., 2014a. Classification into homogeneous groups using Combined Cluster and Discriminant Analysis. Environmental Modelling & Software, 57: 52-59.

Kovács, J. et al., 2015a. Spatial optimization of monitoring networks on the examples of a river, a lake-wetland system and a sub-surface water system. Water Resources Management, 29(14): 5275-5294.

*Kovács, J. et al., 2017. The role of annual periodic behavior of water quality parameters in primary production: Chlorophyll-a estimation. Ecological Indicators, 78C: 311-321.

*Kovács, J., Hatvani, I.G., Korponai, J., Kovács, I.Sz., 2010. Morlet wavelet and autocorrelation analysis of long-term data series of the Kis-Balaton Water Protection System (KBWPS). Ecological Engineering, 36(10): 1469-1477.

*Kovács, J., Korponai, J., Székely Kovács, I., Hatvani, I.G., 2012a. Introducing sampling frequency estimation using variograms in water research with the example of nutrient loads in the Kis-Balaton Water Protection System (W Hungary). Ecological Engineering, 42: 237-243.

*Magyar, N. et al., 2013. Application of multivariate statistical methods in determining spatial changes in water quality in the Austrian part of Neusiedler See. Ecological Engineering, 55: 82-92.

*Magyar, N. et al., 2021. A new approach in determining the decadal common trends in the groundwater table of the watershed of lake "Neusiedlersee". Water, 13(3): 290.

*Tanos, P., Kovács, J., Kovács, S., Anda, A., Hatvani, I., 2015. Optimization of the monitoring network on the River Tisza (Central Europe, Hungary) using Combined Cluster and Discriminant Analysis, taking seasonality into account. Environmental Monitoring and Assessment, 187(9): 575 1-14.

További hivatkozott irodalmak

Abbott, B.W. et al., 2019. Human domination of the global water cycle absent from depictions and perceptions. Nature Geoscience, 12(7): 533-540.

Abdulaziz, A. et al., 2023. The distribution of fecal contamination in an urbanized tropical lake and incidence of acute diarrheal disease. ACS ES & T Water, 3(6), 1561-1573.

Aggarwal, P.K., Gat, J.R., Froehlich, K.F.O., 2005. Isotopes in the water cycle: Past, present and future of a developing science. Springer Netherlands, Dordrecht, 382 p. ISBN:9781402030239

Anselin, L., 1996. The Moran scatterplot as an ESDA tool to assess local instability in spatial association. In: Fischer, M., Scholten, H.J., Unwin, D. (Eds.), Spatial analytical perspectives on GIS Routledge, London, UK, pp. 111-126. ISBN:0203739051

Antal, E., 1996-2000. Meteorológia. In: Kollega Tarsoly I., György, F. (Eds.), Magyarország a XX. Században. Babits Kiadó, Szekszárd, pp. 369-384.

Antal, E. et al., 1991. Das Klima des Neusiedler Sees. Zentralanstalt für Meteorologie und Geodynamik, Wien - Budapest. ISBN:2310004317850

Araguás-Araguás, L., Froehlich, K., Rozanski, K., 2000. Deuterium and oxygen-18 isotope composition of precipitation and atmospheric moisture. Hydrological Processes, 14(8): 1341-1355.

Bedri, Z. et al., 2016. Evaluating a microbial water quality prediction model for beach management under the revised EU Bathing Water Directive. Journal of Environmental Management, 167: 49-58.

Bindoff, N.L. et al., 2013. Detection and attribution of climate change: From global to regional. In: Stocker, T.F. et al. (Eds.), Climate change 2013: The physical science basis. IPCC working group i contribution to ar5. Cambridge University Press, Cambridge, UK, pp. 867-952.

Bíró, G., Braun, G., Oldal, I., 1981. Environmental-sanitary bacteriological studies on the southern watershed area of Lake Balaton. Egészségtudomány, 25(1): 65-72.

Blöschl, G. et al., 2019. Twenty-three unsolved problems in hydrology (uph) – a community perspective. Hydrological Sciences Journal, 64(10): 1141-1158.

Boumaiza, L., Chesnaux, R., Drias, T., Walter, J., Stumpp, C., 2021. Using vadose-zone water stable isotope profiles for assessing groundwater recharge under different climatic conditions. Hydrological Sciences Journal, 66(10): 1597-1609.

Bowen, G.J., 2008. Spatial analysis of the intra-annual variation of precipitation isotope ratios and its climatological corollaries. Journal of Geophysical Research: Atmospheres, 113(D5).

Bowen, G.J., 2010. Isoscapes: Spatial pattern in isotopic biogeochemistry. Annual Review of Earth and Planetary Sciences, 38(1): 161-187.

Bowen, G.J., Cai, Z., Fiorella, R.P., Putman, A.L., 2019. Isotopes in the water cycle: Regional- to global-scale patterns and applications. Annual Review of Earth and Planetary Sciences, 47(1): 453-479

Brkić, Ž., Kuhta, M., Hunjak, T., Larva, O., 2020. Regional isotopic signatures of groundwater in Croatia. Water, 12(7): 1983.

Burt, T.P., Howden, N.J.K., Worrall, F., 2014. On the importance of very long-term water quality records. WIREs Water, 1(1): 41-48.

Çamdevýren, H., Demýr, N., Kanik, A., Keskýn, S., 2005. Use of principal component scores in multiple linear regression models for prediction of chlorophyll-a in reservoirs. Ecological Modelling, 181(4): 581-589.

Carpenter, S.R., 2008. Phosphorus control is critical to mitigating eutrophication. Proceedings of the National Academy of Sciences, 105(32): 11039-11040.

Cattell, R.B., 1966. The scree test for the number of factors. Multivariate Behavioral Research, 1(2): 245-276.

Celle-Jeanton, H., Travi, Y., Blavoux, B., 2001. Isotopic typology of the precipitation in the Western Mediterranean region at three different time scales. Geophysical Research Letters, 28(7): 1215-1218.

Chapman, D.V., 1996. Water quality assessments: A guide to the use of biota, sediments, and water in environmental monitoring. E & FN Spon, London, 626 pp. ISBN:0419216006

Chilès, J.-P., Delfiner, P., 2012. Geostatistics. Wiley, Canada, 731 pp. ISBN:9780470183151

Christofi, C., Bruggeman, A., Kuells, C., Constantinou, C., 2020. Isotope hydrology and hydrogeochemical modeling of troodos fractured aquifer, Cyprus: The development of hydrogeological descriptions of observed water types. Applied Geochemistry, 123: 104780.

Clark, I.D., Fritz, P., 1997. Environmental isotopes in hydrogeology. Taylor & Francis, 342 p. ISBN:9781566702492

Cooper, C., 2006. Lakes as tourism destination resources. In: C. M. Hall, Haerkoenen, T. (Eds.), Lake tourism. An integrated approach to lacustrine tourism systems. Channel View Publications, Clevendon, Buffalo, Toronto, pp. 27-42. ISBN:1845413776

Coplen, T.B., 1994. Reporting of stable hydrogen, carbon and oxygen isotopic abundances. Pure and Applied Chemistry, 66: 273-276.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Machine Learning, 20(3): 273-297.

Craig, H., 1961. Isotopic variations in meteoric waters. Science, 133(3465): 1702-1703.

Csathó, P. et al., 2007. Agriculture as a source of phosphorus causing eutrophication in Central and Eastern Europe. Soil Use and Management, 23(s1): 36-56.

Dansgaard, W., 1964. Stable isotopes in precipitation. Tellus, 16: 436-468.

Day, W.H.E., Edelsbrunner, H., 1984. Efficient algorithms for agglomerative hierarchical clustering methods. Journal of Classification, 1(1): 7-24.

de Brauwere, A., Ouattara, N.K., Servais, P., 2014. Modeling fecal indicator bacteria concentrations in natural surface waters: A review. Critical Reviews in Environmental Science and Technology, 44(21): 2380-2453.

Demény, A., 2003. Stabilizotóp-geokémia és termometria: Hogyan és mire? Földtani Közlöny 133(2): 263-270.

Dinka, M., Kiss, A., Magyar, N., Ágoston-Szabó, E., 2016. Effects of the introduction of pre-treated wastewater in a shallow lake reed stand. Open Geosciences, 8(1): 62-77.

Dinka, M., Szeglet, P., 1998. Reed (*Phragmites australis* (cav.) Trin. ex Steudel) growth and production in different habitats of Neusiedlersee (Lake Fertö). SIL Proceedings 1922-2010., 26(4): 1830-1834.

Dokulil, M.T., 2014. Environmental impacts of tourism on lakes. In: Ansari, A.A., Gill, S.S. (Eds.), Eutrophication: Causes, consequences and control: Volume 2. Springer Netherlands, Dordrecht, pp. 81-88. ISBN:9789400778146

Donohue, I., Styles, D., Coxon, C., Irvine, K., 2005. Importance of spatial and temporal patterns for assessment of risk of diffuse nutrient emissions to surface waters. Journal of Hydrology, 304(1): 183-192.

Döll, P., Fiedler, K., Zhang, J., 2009. Global-scale analysis of river flow alterations due to water withdrawals and reservoirs. Hydrology and Earth System Sciences, 13(12): 2413-2432.

Döll, P. et al., 2012. Impact of water withdrawals from groundwater and surface water on continental water storage variations. Journal of Geodynamics, 59-60: 143-156.

Döll, P., Siebert, S., 2002. Global modeling of irrigation water requirements. Water Resources Research, 38(4): 8.1-8.10.

Driver, H.E., Kroeber, A.L., 1932. Quantitative expression of cultural relationships. University of California Publications in American Archaeology and Ethnology, 31(4): 211-256.

Duda, R.O., Hart, P.E., Stork, D.G., 2012. Pattern classification. John Wiley & Sons, New York, USA, 688 p. ISBN:111858600X

EC, 2000. Directive 2000/60/european commission of the european parliament and of the council of 23 october 2000 establishing a framework for community action in the field of water policy. Official Journal of the European Communities, L 327: 1-72.

EC, 2006a. Directive 2006/7/ec of the european parliament and of the council of 15 february 2006 concerning the management of bathing water quality and repealing directive 76/160/eec. Official Journal of the European Communities, 64: 1-37.

EC, 2006b. Directive 2006/118/EC of the European Parliament and of the Council on the protection of groundwrater against pollution and deterioration. In: European Commission (Ed.), Brussels, pp. 1-72.

EC, 2007. Pursuant to article 16 of regulation (ec) no 648/2004 of the European Parliament and of the Council of 31 March 2004 on detergents, concerning the use of phosphates. In: European Union (Ed.), Brussels.

EEA, 2023. State of bathing waters, European Environment Agency, Brussles. https://www.eea.europa.eu/themes/water/interactive/bathing/state-of-bathing-waters Utóljára megtekintve 2023.12.20.

El Ouali, A. et al., 2022. Isotopic characterization of rainwater for the development of a local meteoric water line in an arid climate: The case of the Wadi Ziz Watershed (South-Eastern Morocco). Water, 14(5): 779.

ElGhawi, R., Pekhazis, K., Doummar, J., 2021. Multi-regression analysis between stable isotope composition and hydrochemical parameters in karst springs to provide insights into groundwater origin and subsurface processes: Regional application to Lebanon. Environmental Earth Sciences, 80(6): 230.

Faragó, T., 2014. Világtalálkozók a vízről (World Summits on Water). Fenntartható Fejlődés Plusz, 3: 10-13.

Fink, G., Alcamo, J., Flörke, M., Reder, K., 2018. Phosphorus loadings to the world's largest lakes: Sources and trends. Global Biogeochemical Cycles, 32(4): 617-634.

Fisher, R.A., 1936. The use of multiple measurements in taxonomic problems. Annals of Eugenics, 7(2): 179-188.

Fórizs, I., 2005. Processes behind the isotopic water line: Water cycle and climate. Studia Universitatis Babes-Bolyai, Physica, 50(3): 138-146.

Freeden, W., 2010. Geomathematics: Its Role, Its Aim, and Its Potential. In: Freeden, W., Nashed, M.Z., Sonar, T. (Eds.), Handbook of Geomathematics. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1-42. ISBN:978-3-642-01546-5

Füst, A. és Geiger, J., 2010. Monitoringtervezés és -értékelés geostatisztikai módszerekkel I. Szakértői véleményen alapuló, „igazoló” mintázás geostatisztikai támogatása. Földtani Közlöny, 140: 303-312.

García-Ruiz, J.M., López-Moreno, J.I., Vicente-Serrano, S.M., Lasanta–Martínez, T., Beguería, S., 2011. Mediterranean water resources in a global change scenario. Earth-Science Reviews, 105(3): 121-139.

Gat, J.R., 2005. Some classical concepts of isotope hydrology. In: Aggarwal, P.K., Gat, J.R., Froehlich, K.F.O. (Eds.), Isotopes in the water cycle: Past, present and future of a developing science. Springer Netherlands, Dordrecht, pp. 127-137. ISBN:9781402030239

Gat, J.R., Dansgaard, W., 1972. Stable isotope survey of the fresh water occurrences in Israel and the Northern Jordan Rift Valley. Journal of Hydrology, 16(3): 177-211.

Gat, J.R. et al., 2003. Isotope composition of air moisture over the Mediterranean Sea: An index of the air-sea interaction pattern. Tellus B: Chemical and Physical Meteorology, 55(5): 953-965.

Gat, J.R., Mook, W.G., Meijer, H.A., 2001. Environmental isotopes in the hydrological cycle, 2. International Atomic Energy Agency, Paris, 73 pp.

Geweke, J.F., 1977. The dynamic factor analysis of economic time series models. In: Aigner, D.J., Goldberger, A.S. (Eds.), Latent variables in socio-economic models. North-Holland Pub. Co., Amsterdam: North-Holland, pp. 365-382.

Gosz, J.R., 1993. Ecotone hierarchies. Ecological Applications, 3(3): 369-376.

Goursaud, S., Masson-Delmotte, V., Favier, V., Orsi, A., Werner, M., 2018. Water stable isotope spatio-temporal variability in Antarctica in 1960–2013: observations and simulations from the ECHAM5-wiso atmospheric general circulation model. Climate of the Past, 14(6): 923-946.

Graf, W., Reinwarth, O., Oerter, H., Mayer, C., Lambrecht, A., 1999. Surface accumulation on foundation ice stream, Antarctica. Annals of Glaciology, 29(1): 23-28.

Grinsted, A., Moore, J.C., Jevrejeva, S., 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. Nonlinear Processes Geophysics, 11(5/6): 561-566.

Hammer, C.U., 1989. Dating by physical and chemical seasonal variations and reference horizons. In: Oeschger, H., C.C. Langway, J. (Eds.), The environmental record in glaciers and ice sheets. John Wiley & Sons Limited, pp. 99-121.

Harper, W.V., 2016. Reduced Major Axis Regression. In Wiley StatsRef: Statistics Reference Online (eds N. Balakrishnan, T. Colton, B. Everitt, W. Piegorsch, F. Ruggeri and J.L. Teugels) https://doi.org/10.1002/9781118445112.stat07912 Utóljára megtekintve 2023.12.20.

Hengl, T., Heuvelink, G.B.M., Rossiter, D.G., 2007. About regression-kriging: From equations to case studies. Computers & Geosciences, 33(10): 1301-1315.

Hengl, T., Heuvelink, G.B.M., Stein, A., 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma, 120(1): 75-93.

Herodek, S., 1984. The eutrophication of Lake Balaton: Measurements, modeling and management. Internationale Vereinigung für theoretische und angewandte Limnologie: Verhandlungen, 22(2): 1087-1091.

Herodek, S., Vörös, L., Tóth, F., 1982. A fitoplankton tömege, termelése és a Balaton eutrofizálódása III. Balatonszemesi-medence 1976-1977, Siófoki-medence 1977. Hidrológiai Közlöny, 62(5): 220-229.

Herzfeld, U.C., 2004. Atlas of Antarctica: Topographic maps from geostatistical analysis of satellite radar altimeter data. Springer-Verlag Berlin Heidelberg, Berlin, Germany, 364 p. ISBN:9783540434573

Ho, J.C., Michalak, A.M., Pahlevan, N., 2019. Widespread global increase in intense lake phytoplankton blooms since the 1980s. Nature, 574(7780): 667-670.

Hoerling, M. et al., 2012. On the increased frequency of Mediterranean drought. Journal of Climate, 25(6): 2146-2161.

Hohn, M.E., 1999. Geostatistics and petroleum geology Springer Netherlands, 235 p. ISBN:9789401059015

Honti, M., Gao, C., Istvánovics, V., Clement, A., 2020. Lessons learnt from the long-term management of a large (re)constructed wetland, the Kis-Balaton Protection System (Hungary). Water, 12(3): 659.

Hotelling, H., 1933. Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, 24(6): 417.
Hssaisoune, M. et al., 2022. Isotopic and chemical tracing for residence time and recharge mechanisms of groundwater under semi-arid climate: Case from Rif Mountains (Northern Morocco). Geosciences, 12(2): 74.
IAEA (Nemzetközi Atomenergia Ügynökség), 1992. Statistical treatment of data on environmental isotopes in precipitation. Technical report series International Atomic Energy Agency, Vienna, 781 p. ISBN:9201008929
IAEA, 2005. Isotopic composition of precipitation in the mediterranean basin in relation to air circulation patterns and climate. Tecdoc series. International Atomic Energy Agency, Vienna, Austria, 223 p.
IAEA, 2019. Global Network of Isotopes in Precipitation. The GNIP database. https://nucleus.iaea.org/wiser/index.aspx . Utóljára megtekintve 2023.09.19.
Iglesias, A., Garrote, L., Flores, F., Moneo, M., 2007. Challenges to manage the risk of water scarcity and climate change in the Mediterranean. Water Resources Management, 21(5): 775-788.
Innes, J., 1998. Measuring environmental change. In: Peterson, D.L., Parker, V.T. (Eds.), Ecological scale: Theory and applications. Columbia University Press, New York, New York, NY, USA, pp. 429-457. ISBN:0231105037
IPICS (International Partnerships in Ice Core Sciences), 2006. The IPICS 2k array: A network of ice core climate and climate forcing records for the last two millennia, International Partnerships in Ice Coring Sciences https://pastglobalchanges.org/sites/default/files/download/docs/working_groups/ipics/white-papers/IPICS_2kArray.pdf Utóljára megtekintve 2023.12.21.
Istvánovics, V., 1988. Seasonal variation of phosphorus release from the sediments of shallow Lake Balaton (Hungary). Water Research, 22(12): 1473-1481.
Istvánovics, V., 2009. Eutrophication of lakes and reservoirs. In: Likens, G.E. (Ed.), Encyclopedia of inland waters. Academic Press, Oxford, pp. 157-165. ISBN:9780123706263.
Istvánovics, V. et al., 2007. Updating water quality targets for shallow Lake Balaton (Hungary), recovering from eutrophication. Hydrobiologia, 581: 305-318.
Istvánovics, V., Honti, M., Torma, P., Kousal, J., 2022. Record-setting algal bloom in polymictic Lake Balaton (Hungary): A synergistic impact of climate change and (mis)management. Freshwater Biology, 67(6): 1091-1106.
Istvánovics, V., Osztoics, A., Honti, M., 2004. Dynamics and ecological significance of daily internal load of phosphorus in shallow Lake Balaton, Hungary. Freshwater Biology, 49(3): 232-252.
Izaguirre, I., Allende, L., Romina Schiaffino, M., 2021. Phytoplankton in Antarctic lakes: Biodiversity and main ecological features. Hydrobiologia, 848(1): 177-207.
Jeppesen, E., Søndergaard, M., Meerhoff, M., Lauridsen, T.L., Jensen, J.P., 2007. Shallow lake restoration by nutrient loading reduction—some recent findings and challenges ahead. In: Gulati, R.D., Lammens, E., De Pauw, N., Van Donk, E. (Eds.), Shallow Lakes in a Changing World. Springer Netherlands, Dordrecht, pp. 239-252. ISBN:9789048176267
Jouzel, J., 2013. A brief history of ice core science over the last 50 yr. Climate of the Past, 9(6): 2525-2547.
Józsa, J., Krámer, T., K., H., Napoli, E., Sarkkula, J., 2008. Wind-induced hydrodynamics and sediment transport of Lake Neusiedl - Hungarian-Austrian-Finnish research cooperation from lake-wide to bay-wide scale., Budapest University of Technology

and Economics, Department of Hydraulic and Water Resources Engineering, Budapest.

Kaiser, H.F., 1960. The application of electronic computers to factor analysis. Educational and Psychological Measurement, 20(1): 141-151.

Kardos, M.K., Clement, A., 2020. Similarities among small watercourses based on multiparameter physico-chemical measurements. Central European Geology, 63(1): 27-37.

Kendall, C., Coplen, T.B., 2001. Distribution of oxygen-18 and deuterium in river waters across the United States. Hydrological Processes, 15(7): 1363-1393.

Kisekka, I., Migliaccio, K.W., Muñoz-Carpena, R., Schaffer, B., Li, Y.C., 2013. Dynamic factor analysis of surface water management impacts on soil and bedrock water contents in southern florida lowlands. Journal of Hydrology, 488: 55-72.

Kirschner, A., Herzig, A., Palffy, K., Wolfram, G., 2014. Mikrobiologie und Hygiene. In: Wolfram, G., Déri, L., Zech, S. (Eds.), Strategic study for Lake Neusiedler See - phase 1. Austrian-Hungarian Water Commission, Vienna - Szombathely, 246 p.

Koeniger, P., Margane, A., Abi-Rizk, J., Himmelsbach, T., 2017. Stable isotope-based mean catchment altitudes of springs in the Lebanon Mountains. Hydrological Processes, 31(21): 3708-3718.

Kolander, R., Tylkowski, J., 2008. Hydrochemical seasons in the Lake Gardno catchment on Wolin Island (north-western poland). Limnological Review, 8(1-2): 27-34.

Konecky, B., McKay, N., 2020. NOAA/WDS paleoclimatology - iso2k database global common era paleo-δ^{18}O and δ^{2}H records, NOAA National Centers for Environmental Information.

Konecky, B.L. et al., 2020. The iso2k database: A global compilation of paleo-δ^{18}O and δ^{2}H records to aid understanding of common era climate. Earth System Science Data, 12(3): 2261-2288.

Kongprajug, A. et al., 2021. Integrated analyses of fecal indicator bacteria, microbial source tracking markers, and pathogens for Southeast Asian beach water quality assessment. Water Research, 203: 117479.

Korponai, J. et al., 2010. Transition from shallow lake to a wetland: A multi-proxy case study in Zalavári Pond, Lake Balaton, Hungary. Hydrobiologia, 641(1): 225-244.

Kottek, M.G., Jürgen; Beck, Christoph; Rudolf, Bruno; Rubel, Franz, 2006. World Map of the Köppen-Geiger climate classification updated. Meteorologische Zeitschrift, 15(3): 259-263.

Koutsoyiannis, D., 2020. Revisiting the global hydrological cycle: Is it intensifying? Hydrology and Earth System Sciences 24(8): 3899-3932.

Kovács, J., 2018. Néhány adatelemző módszer alkalmazása földtudományi feladatok megoldására, különös tekintettel a csoportosító eljárásokra, Budapest, MTA nagydoktori disszertáció.

Kovács, J., Márkus, L., Szalai, J., Kovács, I.Sz., 2015b. Detection and evaluation of changes induced by the diversion of River Danube in the territorial appearance of latent effects governing shallow-groundwater fluctuations. Journal of Hydrology, 520: 314-325.

Kovács, J., Márkus, L., Halupka, G., 2004. Dynamic factor analysis for quantifying aquifer vulnerability. Acta Geologica Hungarica, 47(1): 1-17.

Kovács, S., Kovács, J., Tanos, P., 2014b. Combined cluster and discriminant analysis: Package ‘`ccda`’ in `R`

Kubu, G., Krámer, T., Szilágyi, J., 2014. Hydrology and water management. In: Wolfram, G., Déri, L., Zech, S. (Eds.), Strategic study for Lake Neusiedler See - phase 1. Austrian-Hungarian Water Commission, Vienna - Szombathely, 246 p.

Kuramoto, T. et al., 2011. Seasonal variations of snow chemistry at NEEM, Greenland. Annals of Glaciology, 52(58): 193-200.

Lachniet, M.S., Lawson, D.E., Stephen, H., Sloat, A.R., Patterson, W.P., 2016. Isoscapes of δ^{18}O and δ^{2}H reveal climatic forcings on Alaska and Yukon precipitation. Water Resources Research, 52(8): 6575-6586.

Lau, S.S.S., Lane, S.N., 2002. Biological and chemical factors influencing shallow lake eutrophication: A long-term study. Science of The Total Environment, 288(3): 167-181.

Leboulanger, C. et al., 2021. Urban inputs of fecal bacteria to the coastal zone of libreville, gabon, central Western Africa. Marine Pollution Bulletin, 168: 112478.

Lécuyer, C., Bojar, A.-V., Daux, V., Legendre, S., 2020. Geographic variations in the slope of the δ^{2}H– δ^{18}O meteoric water line over Europe: A record of increasing continentality. Geological Society, London, Special Publications, 507: SP507-2020-68.

Leeming, R., Latham, V., Rayner, M., Nichols, P., 1997. Detecting and distinguishing sources of sewage pollution in Australian inland and coastal waters and sediments. In: Eganhouse, R.P. (Ed.), Molecular markers in environmental geochemistry. Acs symposium series. American Chemical Society, pp. 306-319. ISBN:9780841235182

Li, J., Heap, A.D., Potter, A., Daniell, J.J., 2011. Application of machine learning methods to spatial interpolation of environmental variables. Environmental Modelling & Software, 26(12): 1647-1659.

Liotta, M., Favara, R., Valenza, M., 2006. Isotopic composition of the precipitations in the Central Mediterranean: Origin marks and orographic precipitation effects. Journal of Geophysical Research: Atmospheres, 111(D19).

Liotta, M. et al., 2013. Isotopic composition of precipitation and groundwater in Sicily, Italy. Applied Geochemistry, 34: 199-206.

Lis, G., Wassenaar, L.I., Hendry, M.J., 2008. High-precision laser spectroscopy D/H and ^{18}O/^{16}O measurements of microliter natural water samples. Analytical Chemistry, 80(1): 287-293.

Litke, D.W., 1999. Review of phosphorus control measures in the united states and their effects on water quality, Water-Resources Investigations Report 99–4007. US Department of the Interior, US Geological Survey, Denver Colorado USA 38p.

Liu, Y., Wang, Y., Zhang, J., 2012. New machine learning algorithm: Random forest. In: Liu, B., Ma, M., Chang, J. (Eds.), Information Computing and Applications. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 246-252.

López-López, E., Sedeño-Díaz, J.E., Ruiz-Picos, R.A., 2016. Shallow lakes of the Mexican Central Plateau: Assessing their health condition with oxidative stress biomarkers in sentinel organisms. In: Rashed, M.N. (Ed.), Lake sciences and climate change. Intech. InTech, pp. 63-79. ISBN:9789535125570

Lorius, C., Merlivat, L., 1977. Distribution of mean surface stable isotope values in East Antarctica: Observed changes with depth in the coastal area. International Association of Hydrological Sciences Publication, 118: 127-137.

Lotz, G., 1988. A Kis-Balaton Vízvédelmi Rendszer. Hidrologiai Tájékoztató, 10: 20-22.

Löffler, H., 1979. Neusiedlersee: The limnology of a shallow lake in Central Europe. Dr. W. Junk bv Publishers, De Hague. ISBN:9400991681

Macias-Fauria, M., Grinsted, A., Helama, S., Holopainen, J., 2012. Persistence matters: Estimation of the statistical significance of paleoclimatic reconstruction statistics from autocorrelated time series. Dendrochronologia, 30(2): 179-187.

Markad, A.T., Landge, A.T., Nayak, B.B., Inamdar, A.B., Mishra, A.K., 2019. Trophic state modeling for shallow freshwater reservoir: A new approach. Environmental Monitoring and Assessment, 191(9): 586.

Márkus, L., Berke, O., Kovács, J., Urfer, W., 1999. Spatial prediction of the intensity of latent effects governing hydrogeological phenomena. Environmetrics, 10(5): 633-654.

Masson-Delmotte, V. et al., 2008. A review of Antarctic surface snow isotopic composition: Observations, atmospheric circulation, and isotopic modeling. Journal of Climate, 21(13): 3359-3387

Matheron, G., 1963. Principles of geostatistics. Economic geology, 58: 1246-1266.

Mekonnen, M.M., Hoekstra, A.Y., 2016. Four billion people facing severe water scarcity. Science Advances, 2(2): e1500323.

Merlivat, L., Jouzel, J., 1979. Global climatic interpretation of the deuterium-oxygen 18 relationship for precipitation. Journal of Geophysical Research: Oceans, 84(C8): 5029-5033.

Moran, P.A.P., 1948. The interpretation of statistical maps. Journal of the Royal Statistical Society. Series B (Methodological), 10(2): 243-251.

Muli, J.R., 1996. Environmental problems of lake Victoria (East Africa): What the international community can do. Lakes & Reservoirs: Research & Management, 2: 47-53.

Muñoz-Carpena, R., Ritter, A., Li, Y.C., 2005. Dynamic factor analysis of groundwater quality trends in an agricultural area adjacent to Everglades National Park. Journal of Contaminant Hydrology, 80(1): 49-70.

Nelson, D.B., Basler, D., Kahmen, A., 2021. Precipitation isotope time series predictions from machine learning applied in Europe. Proceedings of the National Academy of Sciences, 118(26): e2024107118.

Nevers, M.B., Byappanahalli, M.N., Edge, T.A., Whitman, R.L., 2014. Beach science in the Great Lakes. Journal of Great Lakes Research, 40(1): 1-14.

NRC (National Research Council) 1996. Natural climate variability on decade-to-century time scales. National Academies Press, Washington, DC, USA, 644 p. ISBN:030912722X

Nunes, L.M., Cunha, M.C., Ribeiro, L., 2004. Optimal space-time coverage and exploration costs in groundwater monitoring networks. Environ Monit Assess, 93(1): 103-124.

O'Brien, R.M., 2007. A caution regarding rules of thumb for variance inflation factors. Quality & Quantity, 41(5): 673-690.

Oberholster, P.J., Botha, A.-M., Cloete, T.E., 2006. Toxic cyanobacterial blooms in a shallow, artificially mixed urban lake in Colorado, USA. Lakes & Reservoirs: Science, Policy and Management for Sustainable Use, 11(2): 111-123.

Oerter, H. et al., 2000. Accumulation rates in dronning maud land, antarctica, as revealed by dielectric-profiling measurements of shallow firn cores. Annals of Glaciology, 30(1): 27-34.

Olsen, R.L., Chappell, R.W., Loftis, J.C., 2012. Water quality sample collection, data treatment and results presentation for principal components analysis – literature review and Illinois River watershed case study. Water Research, 46(9): 3110-3122.

Ouellet-Bernier, M.-M., de Vernal, A., 2018. Proxy indicators of climate in the past. In: Chiotis, E. (Ed.), Climate changes in the Holocene:. CRC Press, pp. 41-76. ISBN:9781351260244

Padisák, J., 1985. Population dynamics of the freshwater dinoflagellate ceratium hirundinella in the largest shallow lake of Central Europe, lake Balaton, Hungary. Freshwater Biology, 15(1): 43-52

Padisák, J., Reynolds, C.S., 2003. Shallow lakes: The absolute, the relative, the functional and the pragmatic. Hydrobiologia, 506(1): 1-11.
Page, R.M., Lischeid, G., Epting, J., Huggenberger, P., 2012. Principal component analysis of time series for identifying indicator variables for riverine groundwater extraction management. Journal of Hydrology, 432-433: 137-144.
Pannonhalmi, M., 2016. Since 60 years Hungarian-Austrian water commission specific example the cross border Lake Fertő /Neusiedler See. In: Petre Gâştescu, Bretcan, P. (Eds.), 3rd international conference "Water resources and wetlands" Tulcea, Romania, pp. 245-250. ISBN:22857923
Parry, R., 1998. Agricultural phosphorus and water quality: A US Environmental Protection Agency perspective. Journal of Environmental Quality, 27(2): 258-261.
Pomogyi, P., 1996. KIS-BALATON ANKÉT Összefoglaló értékelés a Kis-Balaton Védőrendszer 1991-1995 közötti kutatási eredményeiről. Pannon Agrártudományi Egyetem, Georgikon Mezőgazdaságtudományi Kar.
Poté, J. et al., 2009. Persistence and growth of faecal culturable bacterial indicators in water column and sediments of Vidy Bay, Lake Geneva, Switzerland. Journal of Environmental Sciences, 21(1): 62-69.
Putman, A.L., Fiorella, R.P., Bowen, G.J., Cai, Z., 2019. A global perspective on local meteoric water lines: Meta-analytic insight into fundamental controls and practical constraints. Water Resources Research, 55(8): 6896-6910.
Qin, B., Liu, Z., Havens, K., 2007. Eutrophication of shallow lakes with special reference to Lake Taihu, China, 194. Springer Dordecht, The Netherlands, 328 p. ISBN:1402061587
Read, E.K. et al., 2017. Water quality data for national-scale aquatic research: The water quality portal. Water Resources Research, 53(2): 1735-1745.
Reay, D.S., Nedwell, D.B., Priddle, J., Ellis-Evans, J.C., 1999. Temperature dependence of inorganic nitrogen uptake: Reduced affinity for nitrate at suboptimal temperatures in both algae and bacteria. Applied and Environmental Microbiology, 65(6): 2577-2584.
Robertson, D.M., Perlman, H.A., Narisimhan, T.N., 2022. Hydrological cycle and water budgets. In: Mehner, T., Tockner, K. (Eds.), Encyclopedia of Inland Waters (second edition). Elsevier, Oxford, pp. 19-27. ISBN:9780128220412
Rozanski, K., Araguás-Araguás, L., Gonfiantini, R., 1993. Isotopic patterns in modern global precipitation. In: Swart, P.K., Lohmann, K.C., Mckenzie, J., Savin, S. (Eds.), Climate change in continental isotopic records. American Geophysical Union, USA, pp. 1-36. ISBN:9781118664025
Rozanski, K., Johnsen, S.J., Schotterer, U., Thompson, L.G., 1997. Reconstruction of past climates from stable isotope records of palaeo-precipitation preserved in continental archives. Hydrological Sciences Journal, 42(5): 725-745.
Sagehashi, M., Sakoda, A., Suzuki, M., 2001. A mathematical model of a shallow and eutrophic lake (the Keszthely Basin, Lake Balaton) and simulation of restorative manipulations. Water Research, 35(7): 1675-1686.
Saito, L., Biondi, F., Salas, J.D., Panorska, A.K., Kozubowski, T.J., 2008. A watershed modeling approach to streamflow reconstruction from tree-ring records. Environmental Research Letters, 3(2): 024006.
Santo Domingo, J.W., Bambic, D.G., Edge, T.A., Wuertz, S., 2007. Quo vadis source tracking? Towards a strategic framework for environmental monitoring of fecal pollution. Water Research, 41(16): 3539-3552.
Sas, H., 1990. Lake restoration by reduction of nutrient loading: Expectations, experiences, extrapolations. SIL Proceedings, 1922-2010, 24(1): 247-251.

Schang, C. et al., 2016. Evaluation of techniques for measuring microbial hazards in bathing waters: A comparative study. PLOS ONE, 11(5): e0155848.

Scheffer, M., 2013. Ecology of shallow lakes. Springer Netherlands, Dortdecht, The Netherlands, 357 p. ISBN:9781402031540

Schindler, D.W., 1974. Eutrophication and recovery in experimental lakes: Implications for lake management. Science, 184(4139): 897.

Schindler, D.W., Carpenter, S.R., Chapra, S.C., Hecky, R.E., Orihel, D.M., 2016. Reducing phosphorus to curb lake eutrophication is a success. Environmental Science & Technology, 50(17): 8923-8929.

Schulz, M., Stattegger, K., 1997. Spectrum: Spectral analysis of unevenly spaced paleoclimatic time series. Computers & Geosciences, 23(9): 929-945.

Sebestyén, V. et al., 2017. Aquatic environmental assessment of Lake Balaton in the light of physical-chemical water parameters. Environmental Science and Pollution Research, 24(32): 25355-25371.

Sebestyén, V., Németh, J., Juzsakova, T., Domokos, E., Rédey, Á., 2019. Lake Balaton: Water quality of the largest shallow lake in Central Europe. In: Maurice, P. (Ed.), Encyclopedia of Water. John Wiley & Sons, pp. 1-15. ISBN: 9781119300755

Sen, P.K., 1968. Estimates of the Regression Coefficient Based on Kendall's Tau. Journal of the American Statistical Association, 63(324): 1379-1389.

Serianz, L., Cerar, S., Vreča, P., 2021. Using stable isotopes and major ions to identify recharge characteristics of the alpine groundwater-flow dominated triglavska bistrica river. Geologija, 64(2): 205–220.

Sharp, Z., 2017. Principles of stable isotope geochemistry, 2nd Ed. Prentice Hall, 416 p.

Shiekh, M.A., Jahan, N., Hoque, M.M., 2006. High degree of fecal contamination in river, lake and pond waters in and around Dhaka city of Bangladesh. Pakistan Journal of Biological Sciences, 9(1): 141-144.

Sisák, I., 1993. Mezőgazdasági eredetű nem pontszerű terhelés vizsgálata a tápanyagmérleg módszerével a Balaton nyugati vízgyőjtőjén. Kandidátusi értekezés, Pannon Agrártudományi Egyetem, Georgikon Mezőgazdaságtudományi Kar, Keszthely.

Snyder, W.M., 1962. Some possibilities for multivariate analysis in hydrologic studies. Journal of Geophysical Research (1896-1977), 67(2): 721-729.

Sodemann, H., Stohl, A., 2009. Asymmetries in the moisture origin of Antarctic precipitation. Geophysical Research Letters, 36(22): L22803

Somlyódy, L., 2018. Vízminőség-szabályozás: Fejlődéstörténelem. Hidrológiai Közlöny, 98(2): 5-11.

Somlyódy, L., Herodek, S., Fischer, J., 1983. Eutrophication of shallow lakes. Modeling and management. The Lake Balaton case study; IIASA collaborative Paper, CP-83-S3. International Institute for Applied Systems Analysis, Laxenbourg, Austria, 377 p. ISBN:9783704500656

Somlyódy, L., van Straten, G., 1986. Background to the lake balaton eutrophication problem. In: Somlyódy, L., van Straten, G. (Eds.), Modeling and managing shallow lake eutrophication. Springer, pp. 3-18. ISBN:9783642827075

Sommer, R., Blaschke, A., Farnleitner, A., Kirschner, A., 2018. Nachhaltiges Gesamtkonzept zum Management des mikrobiologisch-hygienischen Zustandes des Neusiedler Sees unter Berücksichtigung der Siedlungswasserwirtschaft der Umlandgemeinden. Phase 1:„Pilotstudie“. Report to the austrian federal ministry of federal ministry of agriculture, environment, forestry and water management, Vienna.

Steig, E.J. et al., 2005. High-resolution ice cores from us itase (West Antarctica): Development and validation of chronologies and determination of precision and accuracy. Annals of Glaciology, 41(1): 77-84.

Steiger, N.J., Steig, E.J., Dee, S.G., Roe, G.H., Hakim, G.J., 2017. Climate reconstruction using data assimilation of water isotope ratios from ice cores. Journal of Geophysical Research: Atmospheres, 122(3): 1545-1568.

Sturm, K., Hoffmann, G., Langmann, B., Stichler, W., 2005. Simulation of δ^{18}O in precipitation by the regional circulation model REMOiso. Hydrological Processes, 19(17): 3425-3444.

Tabachnick, B.G., Fidell, L.S., 2014. Using multivariate statistics. Pearson Education Limited, Boston, MA, USA, 983 p. ISBN:1292021314

Tátrai, I., Mátyás, K., Korponai, J., Paulovits, G., Pomogyi, P., 2000. The role of the Kis-Balaton Water Protection System in the control of water quality of Lake Balaton. Ecological Engineering, 16(1): 73-78.

Terzer-Wassmuth, S., Wassenaar, L.I., Welker, J.M., Araguás-Araguás, L.J., 2021. Improved high-resolution global and regionalized isoscapes of δ^{18}O, δ^{2}H and d-excess in precipitation. Hydrological Processes, 35(6): e14254.

Terzer, S., Wassenaar, L.I., Araguás-Araguás, L.J., Aggarwal, P.K., 2013. Global isoscapes for δ^{18}O and δ^{2}H in precipitation: Improved prediction using regionalized climatic regression models. Hydrology and Earth System Sciences, 17(11): 4713-4728.

Topp, S.N., Pavelsky, T.M., Jensen, D., Simard, M., Ross, M.R.V., 2020. Research trends in the use of remote sensing for inland water quality science: Moving towards multidisciplinary applications. Water, 12(1): 169.

Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. Bulletin of the American Meteorological Society, 79(1): 61-78.

Trabelsi, R., Matsumoto, T., Zouari, K., Trabelsi, M., Kumar, B., 2020. Investigation of paleoclimate signatures in Sfax deep groundwater (Southeastern Tunisia) using environmental isotopes and noble gases. Quaternary International, 547: 208-219.

Trenberth, K.E., Smith, L., Qian, T., Dai, A., Fasullo, J., 2007. Estimates of the global water budget and its annual cycle using observational and model data. Journal of Hydrometeorology, 8(4): 758-769.

Túri, M. et al., 2019. Tracing groundwater recharge conditions based on environmental isotopes and noble gases, Lom Depression, Bulgaria. Journal of Hydrology: Regional Studies, 24: 100611.

Túri, M. et al., 2020. Paleotemperature reconstruction using environmental isotopes and noble gases in groundwater in Morocco. Hydrogeology Journal, 28(3): 973-986.

US Government, 1948. Clean water act: Federal water pollution control act. 33 U.S.C. §1251 et seq., Washington, DC, USA.

Van Buuren, S., 2018. Flexible imputation of missing data. CRC press, Boca Raton, FL, USA, 342 p. ISBN:0429960344

Van der Molen, D.T., Portielje, R., 1999. Multi-lake studies in the Netherlands: Trends in eutrophication. In: Walz, N., Nixdorf, B. (Eds.), Shallow lakes ’98: Trophic interactions in shallow freshwater and brackish waterbodies. Springer Netherlands, Dordrecht, pp. 359-365. ISBN:9789401729864

van der Veer, G., Voerkelius, S., Lorentz, G., Heiss, G., Hoogewerff, J.A., 2009. Spatial interpolation of the deuterium and oxygen-18 composition of global precipitation using temperature as ancillary variable. Journal of Geochemical Exploration, 101(2): 175-184.

Vance, T.R. et al., 2016. Optimal site selection for a high-resolution ice core record in East Antarctica. Clim. Past, 12(3): 595-610.

Vasić, L., Palcsu, L., Fen, H., 2019. Groundwater gravitational circulation of karst Veliko Vrelo and Malo Vrelo springs by isotope and the noble gas method: Case study of the Beljanica Massif. Environmental Earth Sciences, 78(10): 1-7.

Vitousek, P.M., Porder, S., Houlton, B.Z., Chadwick, O.A., 2010. Terrestrial phosphorus limitation: Mechanisms, implications, and nitrogen–phosphorus interactions. Ecological Applications, 20(1): 5-15.
Vollenweider, R.A., Kerekes, J., 1982. Eutrophication of waters. Monitoring, assessment and control. OECD cooperative programme on monitoring of inland waters (eutrophication control). Environment Directorate, OECD, 154 p.
Vörösmarty, C.J. et al., 2010. Global threats to human water security and river biodiversity. Nature, 467(7315): 555-561.
Vörösmarty, C.J., Sahagian, D., 2000. Anthropogenic disturbance of the terrestrial water cycle. BioScience, 50(9): 753-765.
Vreča, P., Kern, Z., 2020. Use of water isotopes in hydrological processes. Water, 12(8): 2227.
Wackernagel, H., 2003. Multivariate Geostatistics. Springer-Verlag Berlin Heidelberg, 387 p. ISBN:9783662052945
Wallis, J.R., 1965. Multivariate statistical methods in hydrology—a comparison using data of known functional relationship. Water Resources Research, 1(4): 447-461.
Wang, Y., Hou, S., Masson-Delmotte, V., Jouzel, J., 2009. A new spatial distribution map of δ^{18}O in Antarctic surface snow. Geophysical Research Letters, 36(6): L06501.
Wang, Y., Hou, S., Masson-Delmotte, V., Jouzel, J., 2010. A generalized additive model for the spatial distribution of stable isotopic composition in antarctic surface snow. Chemical Geology, 271(3–4): 133-141.
Webster, R., Oliver, M.A., 2008. Geostatistics for Environmental Scientists. John Wiley & Sons, Ltd, 330 p. ISBN:9780470517277
Wen, Z. et al., 2019. Quantifying the trophic status of lakes using total light absorption of optically active components. Environmental Pollution, 245: 684-693.
Werner, M., Heimann, M., 2002. Modeling interannual variability of water isotopes in Greenland and Antarctica. Journal of Geophysical Research: Atmospheres, 107(D1): ACL 1-1-ACL 1-13.
Wetzel, R.G., 2001. Limnology. Limnology (third edition). Academic Press, San Diego, 1006 p. ISBN:9780127447605
Wu, P., Christidis, N., Stott, P., 2013. Anthropogenic impact on Earth's hydrological cycle. Nature Climate Change, 3(9): 807-810.
Ziermann, M., Michaletzky, G., 1995. Idősorok faktoranalízise. Szigma, 26(3-4): 77-91.
Zlinszky, A., Mücke, W., Lehner, H., Briese, C., Pfeifer, N., 2012. Categorizing wetland vegetation by airborne laser scanning on Lake Balaton and Kis-Balaton, Hungary. Remote Sensing, 4(6): 1617-1650.

A RÖVID ÉRTEKEZÉS ALAPJÁUL SZOLGÁLÓ KÖZLEMÉNYEK KÜLÖNLENYOMATAI

Journal of Hydrology 521 (2015) 460–469

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Assessing the relationship of background factors governing the water quality of an agricultural watershed with changes in catchment property (W-Hungary)

István Gábor Hatvani [a,b,*], József Kovács [b], László Márkus [c], Adrienne Clement [d], Richárd Hoffmann [e], János Korponai [f,g]

[a] Institute for Geological and Geochemical Research, Research Centre for Astronomy and Earth Sciences, MTA, Budaörsi út 45, H-1112 Budapest, Hungary
[b] Eötvös Loránd University, Department of Physical and Applied Geology, H-1117 Budapest, Pázmány P. stny 1/C, Hungary
[c] Eötvös Loránd University, Department of Probability Theory and Statistics, H-1117 Budapest, Pázmány P. stny 1/C, Hungary
[d] Budapest University of Technology and Economics, Department of Sanitary and Environmental Engineering, H-1111 Budapest, Műegyetem rakpart 3, Hungary
[e] University of Kaposvár, Faculty of Animal Science, Department of Botany and Plant Production, H-7400 Kaposvár, Guba S. u. 40, Hungary
[f] West Transdanubian Water Authority, Department Kis-Balaton, H-8360 Keszthely, Csík Ferenc stny. 1, Hungary
[g] University of West Hungary, Department of Chemistry and Environmental Sciences, H-9700 Szombathely, Károly Gáspár tér 4,Hungary

ARTICLE INFO

Article history:
Received 24 June 2014
Received in revised form 5 November 2014
Accepted 29 November 2014
Available online 6 December 2014
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Wolfgang Nowak, Associate Editor

Keywords:
Agricultural diffuse loads
Brodgar
Dynamic Factor Analysis
Hydrology
Point source nutrient input
River Zala

SUMMARY

With urbanization and the growth of agriculture, the importance of precisely assessing the contribution of diffuse- and point source loads arriving to surface waters is becoming more and more important. Determining their effects, however, is not as straightforward as it at first seems. The main aim of the study was to determine the driving background factors of a river located in an agricultural watershed, and furthermore, to separate the role of the diffuse- and point source nutrient loads. The method used to achieve the aims was Dynamic Factor Analysis. This is an effective tool for exploring time series which describe such phenomena. It is capable of taking into account the lagged correlation structure, thus enabling the researcher to uncover the background processes operating in time series. Dynamic Factor Analysis was applied to the time series (1978–2006) of 21 response parameters measured in the River Zala and 6 explanatory (agricultural, meteorological, water quality, etc.) parameters measured in its watershed.

The results demonstrated that with the aid of Dynamic Factor Analysis the superimposed effects of the socio-economic changes which began in the mid-1980s, and the introduction of advanced wastewater treatment (P removal) in the river catchment in the early 1990s could be separated and their relative importance assessed, as well as that of other determining external factors.

1. Introduction

With urbanization and the growth of agriculture, the importance of precisely assessing diffuse loads arriving to surface waters and determining their effects is becoming more and more important (Albek, 2003), due to the well known fact that in many western European countries, the relative proportion of P load of agricultural origin is high (e.g. Denmark 39%, Finland 79%; Rekolainen et al., 1997).

In the mid-1980s – as a result of the democratic political changes then beginning – the prevailing concept of agricultural development changed considerably (Csathó and Radimszky, 2009, 2012; Herodek et al., 1995) in the countries of the former Soviet Bloc. Instead of total quantity, quality (i.e. exportability), efficiency, and the economic (based on real and exact cost-benefit evaluation), as well as environmental aspects came to the forefront as criteria. According to the new thinking, high state subsidies were withdrawn from mineral fertilizers, while high fertilizer prices, and the unpredictable future of state farms and cooperatives resulted in a drastic reduction of fertilizer usage. Hungarian agriculture and industry were severely restructured around the years of the collapse of the Soviet Union in 1990/91. Fertilizer prices increased more than ten times starting in the 1980s. The fall in fertilizer use was much more dramatic (between ten and twenty times) than in the Western European countries (Csathó et al.,

* Corresponding author at: Institute for Geological and Geochemical Research, Research Centre for Astronomy and Earth Sciences, Hungarian Academy of Sciences, Budaörsi ut 45, H-1112 Budapest, Hungary. Tel.: +36 70317 97 58; fax: +36 1 31 91738.

E-mail addresses: hatvaniig@gmail.com (I.G. Hatvani), kevesolt@geology.elte.hu (J. Kovács), markus@cs.elte.hu (L. Márkus), clement@vkkt.bme.hu (A. Clement), hoffmann.richard@ke.hu (R. Hoffmann), korponai.janos@iif.hu (J. Korponai).

http://dx.doi.org/10.1016/j.jhydrol.2014.11.078
0022-1694/© 2014 Elsevier B.V. All rights reserved.

2007). In 1974–1990 ~10,170 t of P was sold annually in the form of fertilizer in Zala County, while in 1991–2000 this value decreased to ~920 t yr^{-1} (HCSO, 2012).

Parallel to the effects of these socio-economic changes, far-reaching measures were taken to reduce the loads of surface waters such as in the case of the River Zala, the study area of this paper (Fig. 1). As examples the following should be mentioned: installment of waste water treatment plants, installment of sewerage around Lake Balaton (Somlyódy, 1994), as well as actions taken to reduce the diffuse loads originating from arable land, e.g. the banning of liquid manure farms, forest plantation (GOV, 1991), and erosion control measures. The River Zala is the 17th largest river in Hungary based on its average discharge (6 m^3 s^{-1}). It is 123 km long, stretching from the Zala Hills to Lake Balaton. Its catchment area is 2621.8 km^2 upstream of Zalaapáti (Fig. 1), the inlet to the Kis-Balaton Water Protection System (KBWPS) and, in sequence, the inlet to Lake Balaton. It is the largest tributary of Lake Balaton, supplying almost 50% of its total water input and 35–40% of its nutrient input (Lotz, 1988; Kovács et al., 2010); it therefore plays an important role in the water quality management of the lake. To decrease the external nutrient loads brought by the River Zala to Lake Balaton a mitigation wetland (KBWPS) was created at the mouth of the River Zala beneath the study area with respect to water flow (Lotz, 1988; Hatvani et al., 2014). Its purpose was to provide a place for the eutrophication processes of the easternmost basin of Lake Balaton (Kovács et al., 2006). The KBWPS, besides ameliorating the water quality (trophic state) of Lake Balaton, is considered as a unique habitat, rich in flora and fauna; furthermore, it is a highly protected nature conservation area and under the de-jure protection of the Ramsar Convention (1971).

In the river's waters the decrease in nutrients was not exclusively the result of the drop in commercial fertilizer use, but of point source load-reduction measures taken in the watershed as well. In 1991 the new dephosphorization (chemical P precipitation) unit at the Zalaegerszeg (the largest town of the River Zala catchment) Wastewater Treatment Plant was put into operation (Fig. 1), decreasing the quantity and composition of point source nutrient loads reaching the river. These measures, however, concern point loads and not diffuse ones, unsurprisingly, since the handling and estimation of non-point sources is much more difficult (Brigault and Ruban, 2000).

In the Zala catchment it has been demonstrated with very detailed soil analyses and laboratory experiments (on 2168 agricultural lots) that a substantial accumulation of fertilizer phosphorus (and also nitrogen) had taken place in the basin, amounting to roughly 400 kg P ha^{-1} during the last approx. 20 years of intensive fertilizer usage (Németh et al., 1994). This was a kind of "time-bomb" (Isringhausen, 1997), which might be still around and be sustaining the diffusion of diffuse P loads, depending on the current hydrological conditions. Phosphorus of agricultural origin – fertilizers – reaches the waters of the Zala mainly by soil erosion triggered by surface runoff. This phosphorus then becomes bound to soil colloids (Kovács et al., 2012a; Wu et al., 2013; Zuazo et al., 2004). Although phosphorus in sub-surface runoff – more specifically, of shallow groundwater origin – exists, its contribution to P loads in the River Zala is under 5% (Jolánkai and Pintér, 1982). Therefore, its impact on P concentration in the river is minor (Sisák and Pomogyi, 1994). Based on the results obtained from exploring the phosphorus loads of the River Zala, Sisák and Pomogyi (1994) found that surface runoff is present only at more than 5 m^3 s^{-1} river discharge. It functions as a major hydrological pathway for sediment, containing high phosphorus concentrations, and can therefore be considered a major threat to water quality (Ballantine et al., 2009). Naturally, above 5 m^3 s^{-1} suspended solid values show high variability, since the three following processes may occur in parallel: (i) the sedimentation of suspended solids, (ii) the stirring up of the sediment, happening mostly in the case of low water level, and (iii) surface erosion caused by sudden flood events. It should be noted that the load originating from wastewater has a marginal contribution with respect to the other previously mentioned sources, especially during floods. It is also known that besides the drastic drop in fertilizer usage, erosion control measures were taken as well in the late 1980s in Hungary. All of the above mentioned resulted in the decrease of diffuse loads entering the River Zala.

Many models have been developed which work on different spatio-temporal scales to estimate the nutrient loads of surface waters and their watershed (Ouyang, 2012; Sekine et al., 1992), and in particular the River Zala (e.g. Bolla and Kutas, 1984; Clement et al., 1998; Hatvani et al., 2014; Honti et al., 2006; Jolánkai and Bíró, 1999). Nevertheless, the cost-effective management of diffuse P pollution essentially needs watershed-level approximation and model simulations (Campbell et al., 2004). Model simulations are especially necessary if a river monitoring system is underdeveloped or water quality observations are scarce (Kovács et al., 2012a). However, instream measurements can help

Fig. 1. Location of the Zala River catchment and Zalaapáti sampling site representing the waters of the River Zala where it reaches the KBWPS.

to understand the relationship between river water quality and catchment properties and help in the resolution of the difficult issues related to the separation of diffuse and point source nutrient loads (Albek, 2003; Corriveau et al., 2013; Tang et al., 2012). By uncovering these relations, this research can help to achieve the desired aim, the understanding of the role of nutrient contamination of various origin (diffuse and point source), as well as other phenomena, such as the temperature dependence of biological activity, and P removal from wastewater.

It is not our intention here to provide one more model to estimate diffuse nutrient loads arriving to the River Zala, or to improve existing ones; it is to determine which external influences (latent effects) dominate the long-term (1978–2006) behavior of the water quality of a river. The means to achieving this goal is Dynamic Factor Analysis (DFA). The dynamic character is absolutely essential here, since the observations vary over time, and conventional factor analysis is unable to take into account lagged correlations (Geweke, 1977). In other words, the highest correlation between the time series is not present at the same point in time, being instead shifted temporally. The obtained dynamic factors can be correlated to natural phenomena known to influence water quality in general. Beyond this identification, the method also allows us to: (i) quantify the extent of the influence, and (ii) follow its temporal evolution. Thus, the change in the role of diffuse and point source nutrient loads can be recognized.

2. Materials and methods

2.1. Dynamic factor model

In accordance with the goals set, it is not sufficient to use the generally applied deterministic approaches (Wilkinson, 2006), in which nutrient loads are solely described on a physical basis, and consequently compared to changes in water quality. The present paper, therefore, approaches the separation of the phenomena governing the loads, from the side of the water quality parameters through statistics.

The identification of latent effects requires the application of a certain kind of factor analysis modeling. Hence, if a time series structure is present in the explored datasets (as in the case of this study), – as stated before – factor analysis is not suitable in its conventional form for the purposes described in the introduction. *Dynamic* Factor Analysis (DFA) is the proper method to take into account the lagged correlation structure (Geweke, 1977), and provides a much deeper insight into the processes than ordinary deterministic approaches.

DFA modeling is consolidated in the literature into state space representation with common trends; in other words it is designed to identify underlying common trends, background factors or latent effects, especially from multiple time series, while taking into account their mutual relationships as well (Muñoz-Carpena et al., 2005; Ritter and Muñoz-Carpena, 2006).

In the case of the present study, in practice this would mean the aggregation of water quality (response) parameter time series measured in the waters of the River Zala into dynamic factors, which are consequently paired with response parameters such as point source- and diffuse nutrient loads measured in the watershed (explanatory time series).

Numerous studies have been published in recent years in the field of hydrology and hydrogeology in which DFA was successful in providing a solution to the problem of finding and tying background (explanatory) processes to explored time series. For example, Kovács et al. (2004) used the method to quantify aquifer vulnerability, while the method was also used to predict the intensity of latent effects governing changes in groundwater level in karstic (Márkus et al., 1999) and other hydrogeological systems (Kovács et al., 2012b, 2015). In both cases the same approach as the one discussed in the present paper was applied, however for one response variable, though on multiple sampling sites. Ritter and Muñoz-Carpena (2006) also used the time series of one variable (shallow groundwater level) as response parameters at multiple sites in an agricultural area adjacent to the Everglades National Park. Moreover, Muñoz-Carpena et al. (2005) in the same study area applied dynamic factor modeling in order to find common trends in the changes of different groundwater quality variables and tie them to explanatory parameters. Kuo et al. (2013) used the same methodology as Muñoz-Carpena and his colleagues, but for the first time on a riverine system in order to determine the influencing factors of chlorophyll-a concentrations in the Kaoping River Estuary. A further two of the most recent studies that should be mentioned as well are those of Kisekka et al. (2013) and Campo-Bescós et al. (2013). The former remained in the field of hydrology by determining the common trends primarily to explore the background factors driving the temporal variation in soil- and bedrock water content, while in the latter case, only the majority of explanatory parameters were of hydrological origin, while the response one was monthly remote sensing data to determine the vegetation index in a southern African savanna. Moving further away from hydrology, Zuur and Pierce (2004) used the same methodology as those mentioned previously, but on "squid time series", to discover the determining factors in the abundance of the animals. The discussed studies universally underline the wide applicability of the method and encourage its application in the field of hydrology and environmental science.

The DFA model applied in the present study uses an autoregressive structure for factors. It was developed and first published in Ziermann and Michaletzky (1995), and is called E-DFA for short in the present paper. A program written in Fortran was used for the actual calculations. The concept of solution in E-DFA is based on simultaneously minimizing the static and dynamic error terms; that is, errors in state estimation and in prediction. Turning to the details of E-DFA, the basic factor model equation, after Ziermann and Michaletzky (1995), is:

$$Y(t) = \mathbf{A} \cdot F(t) + \varepsilon(t). \tag{1}$$

In Eq. (1) the observations ($Y(t)$; standardized response variables), the factors and the noise $\varepsilon(t)$ have time series structures. The factors $F(t) = (F_1(t), \ldots, F_M(t))'$, $0 \leqslant t \leqslant T$, where T is the last observation time point, are specified to be uncorrelated univariate autoregressive processes of the first order in the present application (Ziermann and Michaletzky, 1995):

$$F_j(t) = c_{j,0} + c_{j,1} \cdot F_j(t-1) + \delta_j(t) \tag{2}$$

where $c_{j,0}, c_{j,1} \in R$ constants, $\delta(t) = (\delta_1(t), \ldots, \delta_M(t))'$ stands for a Gaussian white noise independent of $\varepsilon(t)$. A linear transformation expressed through the non-random and non-time dependent $\mathbf{N} \times \mathbf{M}$ matrix $\mathbf{A}$ creates the observations $Y(t) = (Y_1(t), \ldots, Y_N(t))'$ from the factors $F(t)$ and an additive white noise $\varepsilon(t) = (\varepsilon_1(t), \ldots, \varepsilon_N(t))'$. The problem is that neither the factor nor the noise, but $Y(t)$ is observable. The factors must be estimated from the observations by another linear transformation, **B**. The autoregressive structure of the estimated factors is then described by the autoregressive coefficient matrix, **C**. From the estimated factors the prediction of the observations can be computed by the linear transformation, **D**. The mean squared error arising is the state estimation error obtained as the squared sum of the differences between observations and their predictions from the factors. Another type of error associated with the set-up is the mean squared prediction error of the factors themselves according to the autoregressive structure Eq. (2), called the dynamic error. For a detailed description, see Márkus et al. (1999). The simultaneous minimization of these errors provides the factor solution to (1). Obtaining the exact solution leads to a very complicated optimization problem, except for in very low dimensional cases (Rapcsák, 2002). Therefore algorithmic solutions are needed (Márkus et al., 1999; Ziermann and Michaletzky, 1995) which determine the desired factor solution.

A known alternative to E-DFA is a commercial program called Brodgar (www.brodgar.com). This is described in detail in the publications of Zuur et al. (2003a,b). In the course of the research, it was used only to run pilot analyses on the dataset with **M** common trends + noise using an ***H*** non-diagonal matrix. It gave a smoother factor solution. This might have been the result of the random walk specification, which enforces the use of less variable-generating noise. On the one hand, Brodgar is known to experience difficulties in handling "shorter" time series, resulting in the slight perturbation of the result after each run. This is a consequence of the computation, and does not represent the reliability of the underlying model; nevertheless it is a difficulty one must face when conducting analyses. On the other hand, auto- and partial autocorrelations indicated an autoregressive structure for the observed data, with Akaike Information Criterion (AIC) showing first order. This fact confirms our view that the autoregressive specification follows the fluctuations of the dynamic factors more accurately, and that complies with the E-DFA specification. Therefore E-DFA factors are more suitable for the dataset at hand.

In practice, the basic assumption is that there are common driving forces behind the assessed response parameters, which can be described by trends and/or – as in this case – explanatory variables. The steps of the analysis are the following:

(1) As an initial step DFA is applied to the time series of the response variables, resulting in dynamic factor time series (common trends which describe the latent effects).
(2) The factor loadings of the DF time series are then assessed. This provides information on the intensity of the presence of the response parameters in the common trends/dynamic factors. Those loadings are considered as influential which fall outside the ±0.6 interval. It should be noted that, since the loadings depend on sampling, these are considered to be statistics, random variables. Therefore, a ±0.03 (5%) error is accepted regarding the interval of influence.
(3) As a last step the DF time series are correlated with the explanatory variables' time series to find the corresponding physical governing background process which are actually present and which can be related to the previously obtained factor time series (step 1). Those explanatory parameters are considered as influential which have a near 0.6 or stronger linear relationship with the DF time series.

Regarding software, our own dynamic factor model (E-DFA), CorelDRAW Graphics Suite X6, R 3.1 and MS Excel 2013 were used.

2.2. Dataset used

In the study 21 response parameters were analyzed from the River Zala (1978–2006) at the Zalaapáti sampling site (Fig. 1): Chemical oxygen demand using the potassium-dichromate method (COD_{ps}); biological oxygen demand (BOD_5); Cl^-; SO_4^{2-}; HCO_3^-; CO_3^{2-}; Fe^{2+}; Mn^{2+}; Mg^{2+}; Ca^{2+}; Na^+; K^+; Organic Nitrogen (Org. N); NH_4-N; NO_2-N, NO_3-N; particulate phosphorus (PP); soluble reactive phosphorus (SRP); total suspended solids (TSS, mg l^{-1}); chlorophyll-a (Chl-a; mg m^{-3}); and pH. The samples were taken from the upper one meter water column and analyzed weekly by the laboratory of the West Transdanubian Water Authority's Kis-Balaton Department following the water authority's national code of practice during the time period investigated. In the study latent or background effects governing the variability of the response parameters were sought, and in order to identify these mathematically obtained theoretical latent effects with real life phenomena, explanatory parameters were necessary. These were: annual soil P surplus (kg ha^{-1} y^{-1}) and annual soil N surplus (kg ha^{-1} y^{-1}) calculated for the watershed of the River Zala; annual effluent total P concentration of the Zalaegerszeg WWTP (Pww; mg l^{-1}); water temperature (W_{temp}; °C); annual averages of daily precipitation (1979–2006; mm); and total discharge (Q; m^3 s^{-1}).

Variance inflation was tested between the explanatory variables using the 'usdm' package in R (Getis and Ord, 1996). This highlighted that by excluding N surplus from the set of explanatory parameters, their variance inflation factors all stayed well below the rule-of-thumb, value of 10 (Fig. A1; O'Brian, 2007). Note here that N surplus was only used to provide a broader perspective on the variations of nutrient surplus in the soil; it did not play any role in the statistical analyses.

In order to calculate gross nutrient balances for agricultural fields, manure and fertilizer displacement, crop yield and mineralization effects were taken into account, based on county-level statistical data obtained from the Hungarian Central Statistical Office (HCSO), while Pww concentrations were obtained from the operator of the Zalaegerszeg WWTP. Since exact effluent values were not available before the year 2002, the predetermined threshold limits were used, because the source (the operator at the WWTP) stated that the plant had always aimed to keep the P effluent under and/or at around the predetermined threshold limit. The explanatory parameters describing the watershed were all representative for the Zala catchment. Water discharge, representing total runoff, was measured at the same location as the response parameters, whereas the meteorological parameters were sampled at the Keszthely meteorological station, in the vicinity of the Zalaapáti sampling site.

In order to be able to use the response and explanatory datasets together, their temporal frequency had to be adjusted to the annual frequency of certain explanatory parameters and normalized. This resulted in a dataset containing ~35,000 measured data. Preliminary exploratory data analysis showed that both response and explanatory data possess a lagged correlation, thus a time series structure is present. The correlation matrices of both the explanatory and response parameters can be found in Appendices A2 & A3.

3. Results

3.1. General overview of the response datasets

To give a brief overall picture of the response parameters, their descriptive statistics should be assessed. The statistics (Table 1) concur with the previously discussed characteristics of the river, as in the case of the high variability and maximum concentration of TSS, or the relatively high coefficient of variation (CV) for PP. The reason behind the behavior of both TSS & PP could be extreme flood events, surface erosion, etc. At the same time the concentration of Chl-a is quite low, which is no surprise in such an ecosystem (Hatvani et al., 2011). Chl-a values in rivers are generally lower than in lakes or standing water. In middle-, and even more so in lower-courses, algal biomass (indicated by Chl-a) progresses according to local hydrological conditions and nutrient & light availability (Biggs, 1995; Frankforter et al., 2010; Urrea-Clos et al., 2014). However, in river sections upstream of reservoirs – as in the case of the River Zala – Chl-a content is known to be low, while it increases afterwards due to the changes in the characteristics of the river section caused by the reservoir itself (Ward and Stanford, 2006; e.g. on the River Danube see Liska et al., 2008). Staying with CV, in the case of the 7 major ions (with the exception of carbonate) it is quite low as well, varying between

Table 1
Descriptive statistics of the response parameters; where M = mean, MED = median, SD = standard deviation, CV = coefficient of variation.

Statistic/parameter	M	MED	SD	CV %	MIN	MAX
COD_{ps}	24.33	22.00	8.68	35.66	7.20	82.00
BOD_5	3.73	3.30	2.00	53.55	0.00	15.80
pH	7.98	8.00	0.32	4.00	6.80	8.90
Cl^-	28.64	29.00	7.05	24.62	6.74	65.00
Fe^{2+}	0.49	0.10	1.14	233.50	0.00	13.57
Mn^{2+}	0.13	0.05	0.17	138.16	0.00	1.25
Ca^{2+}	90.77	91.00	13.74	15.14	31.00	134.00
Na^+	29.48	28.00	10.16	34.46	4.10	74.00
K^+	5.55	5.50	1.52	27.48	2.00	13.50
SO_4^{2-}	57.64	54.24	17.60	30.53	17.70	150.00
NH_4-N	0.39	0.12	0.58	146.47	0.00	4.51
Chl-a	6.58	5.00	5.74	87.34	0.00	46.33
Mg^{2+}	33.02	33.00	6.47	19.59	9.70	64.00
HCO_3^-	392.33	403.00	59.62	15.20	97.60	523.00
CO_3^-	2.85	0.00	9.65	338.70	0.00	126.00
TSS	50.39	25.00	113.70	225.63	0.00	1793.00
NO_2-N	0.09	0.04	0.12	139.88	0.00	0.85
NO_3-N	1.95	1.84	0.76	38.98	0.02	7.12
SRP	0.19	0.14	0.18	91.77	0.00	1.50
PP	0.19	0.14	0.20	103.52	0.00	2.19
Org. N	1.45	1.18	1.08	74.84	0.00	9.95

~15.1% and ~34.5% with an average of ~23.9%. The reason behind the outlying behavior of carbonate is the fact that pH changes around 8, which is quite close to the point where carbonate can be present in surface waters.

Taking an overview of the events of the mid-‘80s and early ‘90s, a major change was suspected in the time series of the response parameters. The grounds for this suspicion became evident with the exploration of the time series (Fig. 2), and the change was presumably caused by the previously discussed phenomena, ending with the introduction of P removal at the Zalaegerszeg WWTP. The first and most explicit peak in TSS was observed in 1985. It turned out there were two extreme flood events in May and July of that particular year (see Table 1, TSS maximum). These two events raised the TSS and PP concentration in the water. This evolved in the annual averages as well. This peak in 1985 and the measures taken resulted in the decrease of TSS & PP (Fig. 2a). However, the decrease of PP stopped in 1987–1988, settling at around the average characteristic of the ‘70s and ‘80s taken together, only dropping to the current levels after 1991, with the introduction of P removal at the WWTP. This phenomenon was naturally expressed in the decrease in the annual P loads of the river as well (Fig. 2b). As for the other response variables, their standardized values began to decrease around 1988, and the ones related to WW began to fall after 1991, and did not reach their previous concentrations thereafter. Peaks, of course can be observed, as happened in 1999 in the case of TSS, however these are way beneath the pre-1991 peaks. This pattern is true for most of the response variables. Naturally, the concentration of the inorganic parameters may increase from time to time, sometimes reaching the levels which were observed prior to 1991.

3.2. Assessment of the whole time period (1978–2006)

The concept was to explore first the whole time period (1978–2006) and then assess shorter time segments if necessary, that is, if the previously described break in the time series (around 1991) made it necessary. The determination of the AIC and Bayesian information criterion (BIC) of the factor models of different order indicated that either two or three factors should be accounted for (Table 2). But since, it is known that AIC tends to over-, while BIC tends to underestimate the correct order, the root-mean-square error (RMSE) had to be calculated. This facilitated the choice of three factors for further analysis.

Of the three modeled factors, only the first factor showed a relationship with any of the explanatory parameters. The correlation coefficients between the first factor (E-DF1; Fig. 3) and the P surplus and N-surplus of the soil, and the Pww were 0.85, 0.78 and 0.85 respectively. The figures show similarly decreasing and then stabilizing segments, separated according to the transition point around the late 1980s-early 1990s, when Hungarian industry and agriculture were thoroughly restructured, and the already-mentioned drop in fertilizer usage took place. Furthermore, this was the time when P removal was introduced at WWTPs (1991), resulting in decreased point source loads reaching surface waters.

E-DFA was able to effectively capture the fluctuations in the processes in their entirety, in particular the drop in the P & N surplus (1986–1990; Fig. 3b). The data available for Pww (Fig. 3c) cannot be considered as representing exact measurements prior to 2002 (for details see Section 2.2), and they cannot therefore fully represent small fluctuations taking place beneath the predetermined threshold limits. Nevertheless, since it was only used as an explanatory parameter to see whether its trend follows the DF time series or not, its variance does not play an important role.

Based on the results obtained from the assessment of the whole time interval, the question of whether the background processes change their character over time was raised, as two main trends occur in two different time segments. From 1978–1991 descent dominated, while from 1992 to 2006 a stable fluctuation-dominated process (closer to a stationary regime) can be observed (Fig. 3a). These two time periods are separated by the effects of the socio-economic changes starting in the mid-1980s, and the other water quality amelioration measures discussed earlier. The most explicit and important change was seen in 1991, when the previously decreasing trends bottomed out and P removal was introduced at the WWTP of Zalaegerszeg. This was an important event, since P is the most critical parameter in the protection of Lake Balaton and its whole watershed (Istvánovics et al., 2007). The impact of these phenomena changed the behavior of the river's processes to such an extent that the analysis of the entire time series cannot provide coherent results; therefore, the separation of the dataset at the breakpoint becomes a necessity.

3.3. Assessment of the split time series (1978–1991 & 1992–2006)

After splitting the time series in two, the factor loadings of the parameters were analyzed. Between 1978 and 1991 there was no parameter which had an influential factor loading in the first DF time series (DF1). Their factor loadings were more or less evenly distributed, meaning that there is no significant difference between the loadings (M = 0.007; MIN = −0.242; MAX = 0.323). In the second factor (DF2) Mn^{2+}, NH_4-N and PP, while in the third (DF3) COD_{ps}, SO_4^{2-}, TSS and PP had considerable factor loadings (Table 3). Exploring the factor loadings (response parameters) alone is not enough to obtain sufficient information about the background processes; for this, the connection of the factors and the explanatory parameters also has to be assessed. With regard to explanatory parameters, DF1 time series correlated with precipitation ($r = 0.77$; Fig. 4a), and DF2 with the P surplus of the soil ($r = 0.67$) and the Pww ($r = 0.59$; Fig. 4b).

In the second time period (1992–2006) in the first DF, ions and PP had influential loadings, in the second DF parameters in close connection to organic matter content and the redox conditions had the most influential factor loadings, while in the third the most influential loadings were those relating to Org. N, and COD_{ps} (Table 4).

A more detailed analysis of the factor time series and the explanatory variables revealed that the first factor is closely related

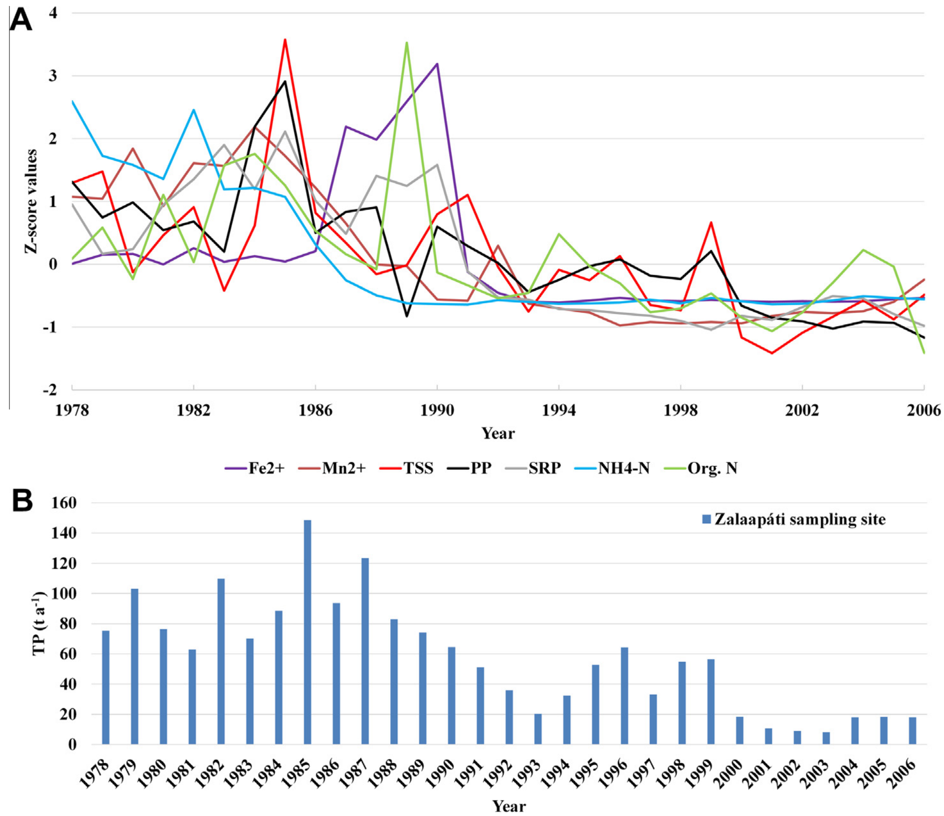


Fig. 2. Z-score (standardized) values of response parameters: Fe^{2+}; Mn^{2+}; TSS; PP; SRP; NH_4-N; Org. N (A) and phosphorus loads (B) at the Zalaapáti sampling site for 1978–2006.

Table 2
Model statistics of the different order with the minimum AIC, BIC and the smallest RMSE underlined.

Factor	AIC	BIC	RMSE
1	−1174.2	−1065.2	0.373
2	−3025.8	−2873.8	0.302
3	−3091.7	−2698.8	0.298
4	−2990.5	−2554.6	0.301
5	−2441.6	−1896.7	0.432
6	−2393.7	−1739.7	0.432

to the discharge of the River Zala ($r = 0.74$), and to the Pww ($r = 0.65$; Fig. 5a). Despite the close linear connection to Pww, DF1 shows no significant correlation whatsoever with the soil P surplus ($r = -0.076$). Meanwhile, the third factor corresponds to water temperature ($r = 0.77$; Fig. 5b).

4. Discussion

It should be stated that although significant linear relations were observed between the DF time series and the soil P surplus, along with the P output of the Zalaegerszeg WWTP, it is generally difficult to separate their effect. What does become clear, however, from the results is that there is a transition in the relationship of the DF and both the response and explanatory parameters, with a "border" at the year 1991. In the first time period, the DF time series correspond to both the diffuse and point source nutrient loads of the River Zala, while after 1991 this changes, and the role of the P originating from the WWTP increases, along with the discharge of the river and other parameters (Fig. A4).

In the early years of the period 1978–1991 agricultural production was at its peak in almost all the countries of the Soviet Bloc, and Hungary was no exception. The use of fertilizers in agriculture boomed in the beginning of this period thanks to substantial subsidies and the highly industrialized development of fertilizer production. Precipitation generates surface runoff and soil erosion, which are the main transport mechanisms in the catchment (Albek, 2003). In the case of flood events and/or extreme weather conditions P compounds associated with soil particles are transported into the river, causing diffuse nutrient loads. Besides these diffuse loads, there was a significant point source as well the effluent of the Zalaegerszeg WWTP. This may be the explanation behind the fact that precipitation, fertilizer inputs (P surplus in the soil originating explicitly from agriculture), Pww, and other related parameters (NH_4-N, PP etc.) were represented in the DF time series.

The phenomenon observed, that there were no influential response water quality parameters in DF1 may have as its cause the fact that in 1978–1991 the external influences reaching the watershed were mainly driven by precipitation. Therefore, in this particular period the response variables were influenced in a similar and overwhelming way. Other specific effects concerning the response parameters are still present, but expressed in the other factors. Specifically, PP indicates the relationship of DF2 to agriculture, to Pww, and, presumably, to the stirring up of the sediment, while NH_4-N represents only cleaned WW inputs to the river. As for DF3, although there is no explanatory parameter related to the DF in question, the influential factor loading of TSS and PP can be related to erosion and sediment washed into the river, specifically diffuse loads.

Between the two analyzed periods (1978–1991 & 1992–2006) a major difference was observed in terms of response and explanatory parameters as well. Between 1992 and 2006 none of the explanatory parameters related to agriculture were influential; however, the relationship between the DF time series and Pww

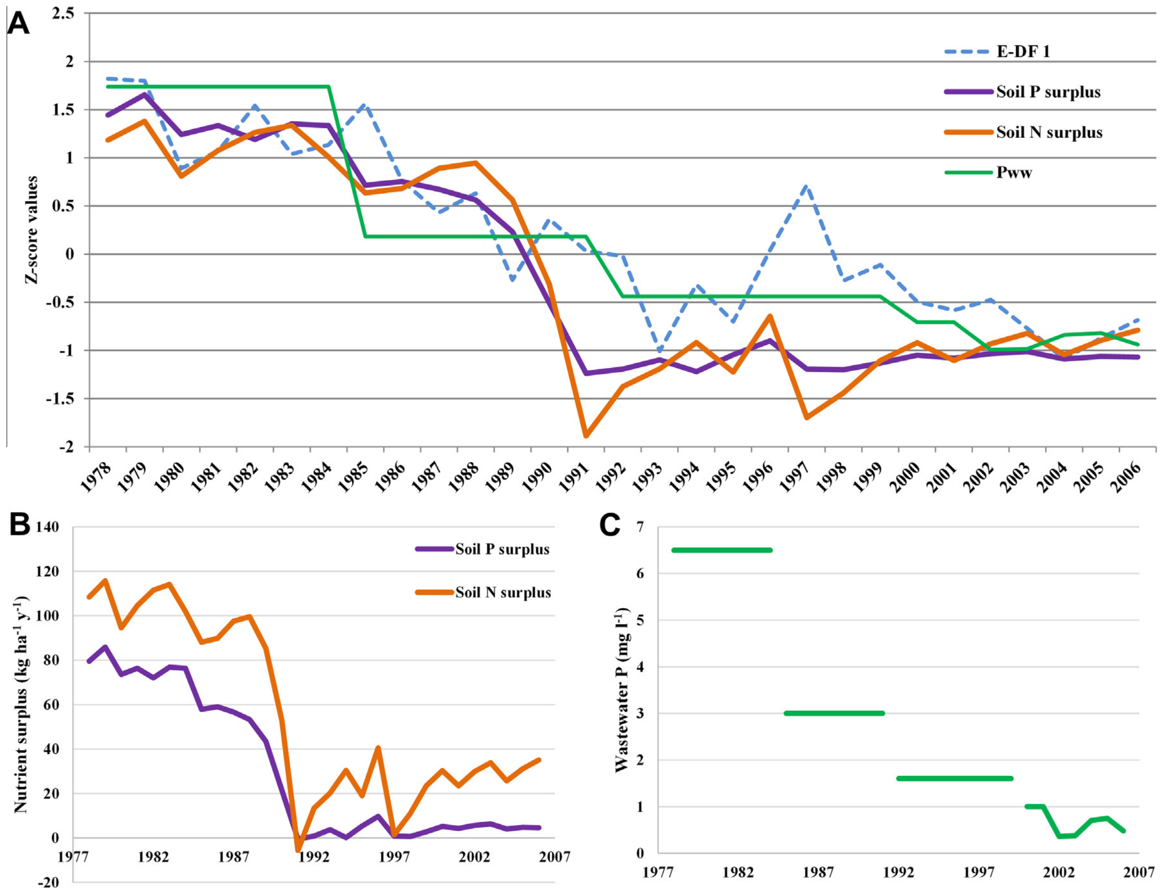


Fig. 3. First dynamic factor time series (E-DF1) visualized together with (A) Z-score standardized soil P & N surplus and the P output of the Zalaegerszeg WWTP, (B) the actual soil N and P surplus data and (C) the actual Pww values.

Table 3
Factor loadings of the response parameters in the first three factors (1978–1991); influential loadings are underlined.

Factor/parameter	DF1	DF2	DF3
COD_{ps}	−0.209	0.395	0.609
BOD_5	−0.131	0.307	0.287
pH	0.158	−0.565	0.083
Cl^-	−0.035	−0.536	0.074
Fe^{2+}	−0.150	−0.365	−0.035
Mn^{2+}	0.093	0.668	−0.028
Ca^{2+}	−0.187	0.415	−0.156
Na^+	−0.015	−0.139	0.130
K^+	−0.242	0.273	0.246
SO_4^{2-}	0.039	0.482	0.571
NH_4-N	0.031	0.627	−0.061
Chl-a	0.001	−0.255	0.169
Mg^{2+}	0.040	−0.382	−0.156
HCO_3^-	−0.014	0.466	−0.521
CO_3^-	0.102	−0.529	0.040
TSS	−0.011	0.257	0.782
NO_2-N	0.183	0.397	−0.268
NO_3-N	0.111	−0.245	0.314
SRP	0.323	0.376	0.236
PP	−0.066	0.593	0.589
Org. N	0.123	0.367	−0.307

remained, and moved up to the first factor, indicating a relatively more important role. Note here, that this does not mean increased concentrations; it means relative importance in relation to the other parameters. The disappearance of the agriculture-related parameters occurred because of the decrease in the use of fertilizers and other chemical compounds related to water quality deterioration (Herodek et al., 1995).

As for the further changes in the stochastic relations of the river's waters detectable by DFA regarding both periods it can be said that since DFA detects the processes and not the concentrations it is obvious that the significant relations between discharge, PP and Pww are due to the increased relative importance of P originating from the Zalaegerszeg WWTP, although it decreased in concentration. Measured PP concentrations in the River Zala partly originated from (i) effluent wastewater, and (ii) partly from the sediment stirred up by sudden flood events ($Q > 5\ m^3\ s^{-1}$). These do contribute to the concentration of P in the River Zala, but only for shorter time intervals (24–36 h), thereby causing short-lived diffuse loads (Sisák and Pomogyi, 1994). Even during periods of low discharge ($Q < 5\ m^3\ s^{-1}$), sediments may act as a storage for P. It should be remarked here that these loads are of internal origin, and less related to agriculture than was the case with DF2 in 1978–1991. When water discharge is low, the residence time of the water increases, and there is more time for sediment–water interactions. Furthermore, the strong linear relationship between Pww and DF1 – and the influential parameters sodium and chloride (associated with loads of wastewater origin) – the increased relative importance of wastewater is underlined caused by the "disappearance" of agriculture-related driving factors. In the meanwhile bicarbonate and sulphate, both main halobic variables in the water, have gained a determining role enforcing the importance of discharge as an explanatory parameters. These phenomena presumably occurred before 1992 as well, but their relative importance was lower.

The phenomena observed in DF2 could be related to the fact that sediment absorbs nutrients (primarily P) based on its Fe^{2+} & Mn^{2+} content, pH, and redox conditions (Eh). Because Eh is

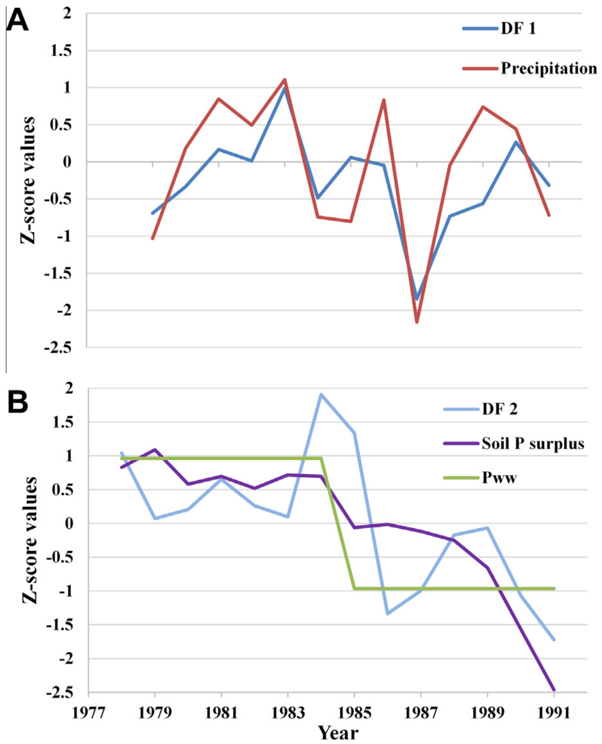


Fig. 4. DF time series visualized together with explanatory parameters: DF1 with precipitation for the time interval 1979–1991 (A) and DF2 with soil P surplus and Pww for the time interval 1978–1991 (B). The explanatory time series were Z-score standardized before plotting.

Table 4
Factor loadings of the response parameters in the first three factors (1992–2006); influential loadings are underlined.

Factor/parameter	DF1	DF2	DF3
COD_{ps}	0.089	0.023	0.729
BOD_5	−0.306	−0.804	−0.122
pH	0.272	0.702	0.136
Cl^-	−0.653	0.325	−0.125
Fe^{2+}	0.313	−0.790	0.072
Mn^{2+}	−0.145	−0.859	−0.110
Ca^{2+}	−0.391	0.367	0.450
Na^+	−0.803	0.260	−0.329
K^+	−0.685	0.103	−0.338
SO_4^{2-}	−0.639	0.132	0.352
NH_4-N	−0.417	−0.437	0.466
Chl-a	0.393	−0.310	−0.442
Mg^{2+}	−0.566	−0.391	0.121
HCO_3^-	−0.710	−0.098	0.070
CO_3^-	0.469	0.318	−0.049
TSS	0.462	−0.369	0.519
NO_2-N	−0.505	0.049	0.107
NO_3-N	−0.454	−0.173	0.525
SRP	−0.492	0.034	−0.179
PP	0.760	0.039	0.082
Org. N	−0.051	0.200	0.587

dependent on pH, a relationship can be inferred between DF2 and the Eh of the water, although these conditions have not been directly measured in the area. Moreover, the opposite signs in the influential factor loadings – in this case – indicate mechanisms operating in opposite directions, i.e. a decrease in pH and Eh would be accompanied by a parallel increase in the solution of Fe^{2+} and Mn^{2+} and the release of P from the sediment, and vice-versa. This means that the loads reaching the River Zala originating from

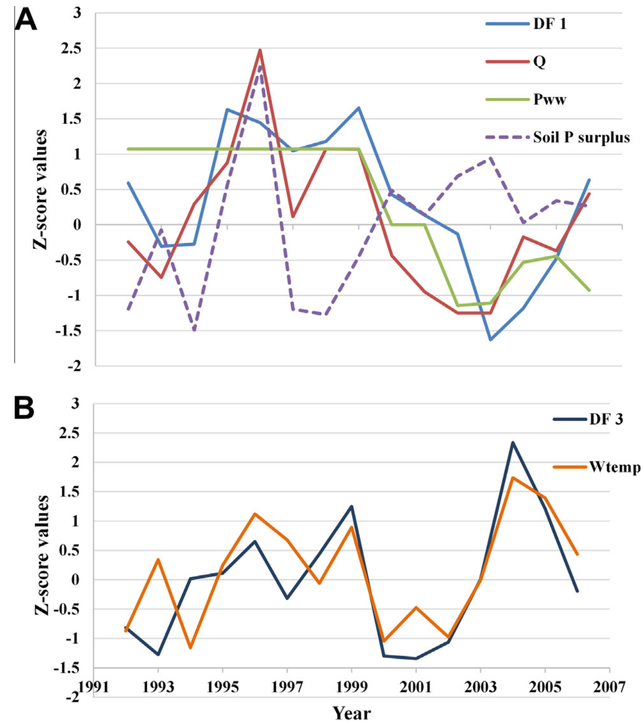


Fig. 5. DF time series visualized together with explanatory parameters: DF1 with discharge and Pww (A) and DF3 with water temperature (B) for the time interval 1992–2006. The explanatory time series were Z-score standardized before plotting.

external sources (e.g. waste water and land use, with decreased importance after 1991, etc.) have "changed places" with the processes that control its internal loads. Internal processes have become more relevant. Phenomena such as the one discussed here also occur nearby in Lake Balaton (Istvánovics et al., 2004).

The temperature-dependence of microbial activity (Van Straten and Herodek, 1982) – reflected in the W_{temp} explanatory parameter – affects the amount of autochthonous organic matter produced in the water. The change in autochthonous organic matter content is appeared in the two influential response parameters in DF3 (Org. N and COD_{ps}). In summary, it can be concluded that the relationship of the response and explanatory parameters to DF3 represent an increased relative driving power of biological activity, which was the result of the relative decrease of external- and increase of internal loads, as seen in DF2.

5. Conclusions

There were three steps in the conduct of the research on the driving factors of the River Zala. First (i) the similarities (dynamic factors/common trends) were extracted from the 21 response time series, then (ii) the loadings of the response parameters were retrieved, describing the intensity of the acting factors and last (iii) the natural phenomena corresponding to the factors were identified. These steps were necessary because the natural processes emerged through a set of such complex transcripts that they cannot be directly determined, and the application of a *factor analysis type of method was therefore needed*. Furthermore, because these natural processes operated through lagged effects, the factor analysis had to be *dynamic*, that is, capable of handling data structured by delayed interdependences such as those in the study.

In conclusion it can be stated that with the application of DFA, the consequences of anthropogenic effects (contamination of waters and/or any other environmental phenomena) can be followed in time, and their governing background processes can be determined and located on a temporal scale. Importantly, from the monitoring of the river's water quality, where the results of all contaminating sources appear superimposed one upon the other, it was possible to separate out the various contaminant sources and follow their temporal evolution. The analysis presented here found what may be the most important governing processes of the water quality of the River Zala: changes in the relative importance of diffuse- and point source nutrient loads, nutrient resuspension from the sediment, temperature dependence of biological activity, etc., along with the external changes in the river catchment, reflected in the DF time series.

Acknowledgements

We the authors would like to thank Paul Thatcher for his work on our English version and say thanks for the stimulating discussions with our colleagues at the Department of Physical and Applied Geology of Eötvös Loránd University. The work of I.G. Hatvani was supported by the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of TÁMOP 4.2.4. A/1-11-1-2012-0001 'National Excellence Program'. In addition, we would like to give thanks for the support of the ''Lendület'' program of the Hungarian Academy of Sciences (LP2012-27/2012). This is contribution No. 21 of 2ka Palæoclimatology Research Group.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jhydrol.2014.11.078.

References

Albek, E., 2003. External phosphorus load estimates and P-budget for the hydroelectric reservoir of Bort-les-Orgues, France. Water Air Soil Pollut. 147, 229–243.
Ballantine, D., Walling, D.E., Leeks, G.J.L., 2009. Mobilization and transport of sediment-associated phosphorus by surface runoff. Water Air Soil Pollut. 196, 311–320.
Biggs, B.J., 1995. The contribution of flood disturbance, catchment geology and land use to the habitat template of periphyton in stream ecosystems. Freshw. Biol. 33, 419–438.
Bolla, M., Kutas, T., 1984. Sub-models for the nutrient loading estimation on River Zala. Ecol. Model. 26, 115–143.
Brigault, S., Ruban, V., 2000. Estimation of point and diffuse contaminant loads to streams by non-parametric regression analysis of monitoring data. Water Air Soil Pollut. 119, 91–103.
Campbell, N., D'Arcy, B., Frost, A., Novotny, V., Sansom, A., 2004. Diffuse Pollution: An Introduction to the Problems and Solutions. IWA Publishing, London.
Campo-Bescós, M.A., Muñoz-Carpena, R., Kaplan, D.A., Southworth, J., Zhu, L., Waylen, P.R., 2013. Beyond precipitation: physiographic gradients dictate the relative importance of environmental drivers on savanna vegetation. PLoS ONE 8 (8), e72348. http://dx.doi.org/10.1371/journal.pone.0072348.
Clement, A., Somlyódy, L., Koncsos, L., 1998. Modeling the phosphorus retention of the Kis-Balaton Upper Reservoir. Wat. Sci. Tech. 37, 113–120.
Corriveau, J., Chambers, P.A., Culp, J.M., 2013. Seasonal variation in nutrient export along streams in the northern great plains. Water Air Soil Pollut. 224, 1594.
Csathó, P., Radimszky, L., 2009. Two worlds within EU27: sharp contrasts in organic and mineral nitrogen–phosphorus use, nitrogen–phosphorus balances, and soil phosphorus status: widening and deepening gap between Western and Central Europe'. Commun. Soil Sci. Plant Anal. 40, 999–1019.
Csathó, P., Radimszky, L., 2012. Sustainable agricultural NP turnover in the EU 27 countries. In: Lichtfouse, E. (Ed.), Organic Fertilization, Soil Quality and Human Health, Sustainable Agricultural Reviews, vol. 9. Springer, Dordrecht-Heidelberg-New York-London, pp. 161–186.
Csathó, P., Sisák, I., Radimszky, L., Lushaj, S., Spiegel, H., Nikolova, M.T., Nikolov, N., Čermák, P., Klir, J., Astover, A., Karklins, A., Lazauskas, S., Kopinski, J., Hera, C., Dumitru, E., Manojlović, M., Bogdanović, D., Torma, S., Leskošek, M. (deceased, 2006), Khristenko, A. 2007. Agriculture as a source of phosphorus causing eutrophication in Central and Eastern Europe. Soil Use and Management 23, 36–56.
Frankforter, J., Weyers, H., Bales, J., Moran, P., Calhoun, D., 2010. The relative influence of nutrients and habitat on stream metabolism in agricultural streams. Environ. Monit. Assess. 168, 461–479.
Getis, A., Ord, J.K., 1996. Local spatial statistics: an overview. In: Longley, P., Batty, M. (Eds.), Spatial Analysis: Modelling in a GIS Environment. Geoinformation International, Cambridge, pp. 261–277.
Geweke, J.F., 1977. The dynamic factor analysis of economic time series models. In: Ainger, D.J., Goldberger, A.S. (Eds.), Latent Variables in Socio-Economic Models. North-Holland, Amsterdam, pp. 365–382.
GOV (Hungarian government), 1991. Governmental edict on forest plantation between 1991–2006. 3224/1991. (VI. 13.).
Hatvani, I.G., Kovács, J., Kovácsné Székely, I., Jakusch, P., Korponai, J., 2011. Analysis of long term water quality changes in the Kis-Balaton Water Protection System with time series-, cluster analysis and Wilks' lambda distribution. Ecol. Eng. 37, 629–635.
Hatvani, I.G., Clement, A., Kovács, J., Kovácsné, Sz.I., Korponai, J., 2014. Assessing water-quality data: the relationship between the water quality amelioration of Lake Balaton and the construction of its mitigation wetland. J. Great Lakes Res. 40, 115–125.
HCSO (Hungarian Central Statistical Office; 2012).
Herodek, S., Istvánovics, V., Jolánkai, G., Csathó, P., Németh, T., Várallyay, Gy., 1995. 1995 P-cycle in the Balaton catchment – a Hungarian case study. In: Phosphorus in the Global Environment. John Wiley and Sons Ltd., pp. 275–300.
Honti, M., Istvánovics, V., Kovács, Á.S., 2006. Balancing between retention and flushing in river networks – optimizing nutrient management to improve trophic state. Sci. Total Environ. 408, 4712–4721.
Isringhausen, S., 1997. GIS-gestützte Prognose und Bilanzirung von Feinboden und Nahrstoffaustragen in einem Teileinzugsgebiet der oberen Lamme in Südniedersachsen Diplomarbeit. Universitat Hannover, pp. 34–42.
Istvánovics, V., Osztoics, A., Honti, M., 2004. Dynamics and ecological significance of daily internal load of phosphorus in shallow Lake Balaton, Hungary. Freshw. Biol. 49, 232–252.
Istvánovics, V., Clement, A., Somlyody, L., Specziar, A., G.-Toth, L., Padisak, J., 2007. Updating water quality targets for shallow Lake Balaton (Hungary), recovering from eutrophication. Hydrobiologia 581, 305–318.
Jolánkai, G., Bíró, P. 1999. Development and application of an integrated catchment model for the River Zala, Hungary. Institute for Water Pollution Control, Budapest (final report) 32p. <http://www.vituki.hu/files/zalacikk_eng.pdf> (Cited on 13.10.2014).
Jolánkai, G., Pintér, Gy., 1982. Areal non point source contamination and surface erosion. Vízügyi Műszaki Gazdasági Tájékoztató 131:155 (in Hungarian).
Kisekka, I., Migliaccio, K.W., Muñoz-Carpena, R., Schaffer, B., Li, Y.C., 2013. J. Hydrol. 488, 55–72.
Kovács, J., Márkus, L., Halupka, G., 2004. Dynamic factor analysis for quantifying aquifer vulnerability. Acta Geol. Hung. 47, 1–17.
Kovács, J., Hatvani, I.G., Korponai, J., Kovácsné, Sz.I., 2006. Morlet wavelet and autocorrelation analysis of long term data series of the Kis-Balaton Water Protection System (KBWPS). Ecol. Eng. 36, 1469–1477.
Kovács, J., Hatvani, I.G., Korponai, J., Kovacsné, Sz.I., 2010. Morlet wavelet and autocorrelation analysis of long term data series of the Kis-Balaton Water Protection System (KBWPS). Ecol. Eng. 36, 1438–1440.
Kovács, A., Honti, M., Zessner, M., Eder, A., Clement, A., Blöschl, G., 2012a. Identification of Phosphorus emission hotspots in agricultural catchments. Sci. Total Environ. 433, 74–88.
Kovács, J., Márkus, L., Szalai, J., Barcza, M., Bernáth, Gy., Kovácsné, Sz.I., Halupka, G., 2012b. In: Voudouris (Ed.), Exploring Potentially Hazardous Areas for Water Quality Using Dynamic Factor Analysis, Water Quality Monitoring and Assessment. InTech. http://dx.doi.org/10.5772/34476, <http://www.intechopen.com/books/water-quality-monitoring-and-assessment/exploring-potentially-hazardous-areas-for-water-quality-by-dinamic-factor-analysis>.
Kovács, J., Márkus, L., Szalai, J., Kovács, Sz.I., 2015. Detection and evaluation of changes induced by the diversion of River Danube in the territorial appearance of latent effects governing shallow-groundwater fluctuations. J. Hydrol. 520, 314–325.
Kuo, Y.-M., Jang, C.-S., Yu, H.-L., Chen, S.-C., Chu, H.-J., 2013. Identifying nearshore groundwater and river hydrochemical variables influencing water quality of Kaoping River Estuary using dynamic factor analysis. J. Hydrol. 486, 39–47.
Liska, I., Wagner, F., Slobodník, J. (Eds.), 2008. Joint Danube Survey 2 – Final Scientific Report. ICPDR – International Commission for the Protection of the Danube River, 242p.
Lotz, Gy., 1988. A Kis-Balaton Vízvédelmi Rendszer (in English: The Kis-Balaton Water Protection System). Hidrol. Tájékoztató. Október 20–22, Hungary.
Márkus, L., Berke, O., Kovács, J., Urfer, W., 1999. Spatial prediction of the intensity of latent effects governing hydrogeological phenomena. Environmetrics 10, 633–654.
Muñoz-Carpena, R., Ritter, A., Li, Y.C., 2005. Dynamic factor analysis of groundwater quality trends in an agricultural area adjacent to Everglades National Park. J. Contam. Hydrol. 80, 49–70.
Németh, T., Csathó, P., Molnár, E., Várallyay, Gy., 1994. Estimation of agronomic NP loads to surface waters in Hungary. In: Vollenbroek, J., Ijjas, I., Bögi, K. (Eds.), Environmental Programme for the Danube River Basin. Haskoning, The Netherlands, pp. 35–61.
O'Brian, R.M., 2007. A caution regarding rules of thumb for variance inflation factors. Qual. Quant. 41, 673–690.
Ouyang, Y., 2012. Estimation of shallow groundwater discharge and nutrient load into a river. Ecol. Eng. 38, 101–104.
Ramsar Convention, ''Convention on Wetlands of International Importance especially as Waterfowl Habitat. Ramsar (Iran), 2 February 1971. UN Treaty

Series No. 14583. As amended by the Paris Protocol, 3 December 1982, and Regina Amendments, 28 May 1987".

Rapcsák, T., 2002. On minimization on Stiefeld Manifolds. Eur. J. Oper. Res. 143, 365–376.

Rekolainen, S., Ekholm, P., Ulén, B., Gustafson, A., 1997. P losses from agriculture to surface waters in the Nordic countries. In: Tunney, H., Carton, O.T., Brookes, P.C., Johnston, A.E. (Eds.), Phosphorus Loss from Soil to Water. CAB Int., London, England, pp. 77–93.

Ritter, A., Muñoz-Carpena, R., 2006. Dynamic factor modeling of ground and surface water levels in an agricultural area adjacent to Everglades National Park. J. Hydrol. 317, 340–354.

Sekine, M., Ukita, M., Nakanishi, H., 1992. Determination of nutrient transformation pathways in rivers using a mathematical programming method. J. Contam. Hydrol. 9, 189–206.

Sisák, I., Pomogyi, P., 1994. Analysis of the loads arriving to the River Zala (in English). Vízügyi Közlemények 76, 417–434.

Somlyódy, L., 1994. Quo vadis water quality management in central and eastern Europe? Water Sci. Technol. 30, 1–14.

Tang, X., Wu, M., Yang, W., Yin, W., Jin, F., Ye, M., Currie, N., Scholz, M., 2012. Ecological strategy for eutrophication control. Water Air Soil Pollut. 223, 723–737.

Urrea-Clos, G., García-Berthou, E., Sabater, S., 2014. Factors explaining the patterns of benthic chlorophyll-a distribution in a large agricultural Iberian watershed (Guadiana river). Ecol. Ind. 36, 463–469.

Van Straten, G., Herodek, S., 1982. Estimation of algal growth parameters from vertical primary production profiles. Ecol. Model. 15, 287–311.

Ward, J.V., Stanford, J.A., 2006. The serial discontinuity concept: extending the model to floodplain rivers. Regul. Rivers: Res. Manage. 10, 159–168.

Wilkinson, D.J., 2006. Stochastic Modelling for Systems Biology. Chapman & Hall/CRC, USA, p. 245.

Wu, M., Tang, X., Li, Q., Yangm, W., Jin, F., Tang, M., Scholz, M., 2013. Review of ecological engineering solutions for rural non-point source water pollution control in Hubei Province, China. Water Air Soil Pollut. 224, 1561.

Ziermann, M., Michaletzky, Gy., 1995. Factor analysis of time series. Sigma J. Soc. Econ. Model. 26, 77–91.

Zuazo, V., Raya, A.M., Ruiz, J.A., 2004. Nutrient losses by runoff and sediment from the taluses of orchard terraces. Water Air Soil Pollut. 153, 355–373.

Zuur, A.F., Pierce, G.J., 2004. Common trends in northeast Atlantic squid time series. J. Sea Res. 52, 57–72.

Zuur, A.F., Fryer, R.J., Jolliffe, I.T., Dekker, R., Beukema, J.J., 2003a. Estimating common trends in multivariate time series using dynamic factor analysis. Environmetrics 14, 665–685.

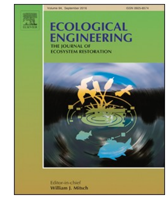
Zuur, A.F., Tuck, I.D., Bailey, N., 2003b. Dynamic factor analysis to estimate common trends in fisheries time series. Can. J. Fish. Aquat. Sci. 60, 542–552.

Ecological Engineering 151 (2020) 105861

Contents lists available at ScienceDirect

Ecological Engineering

journal homepage: www.elsevier.com/locate/ecoleng



Spatiotemporal changes and drivers of trophic status over three decades in the largest shallow lake in Central Europe, Lake Balaton

István Gábor Hatvani[a,*], Vinicius Deganutti de Barros[b], Péter Tanos[c], József Kovács[d], Ilona Székely Kovács[e], Adrienne Clement[f]

[a] Institute for Geological and Geochemical Research, Research Centre for Astronomy and Earth Sciences, Budaörsi út 45, H-1112 Budapest, Hungary
[b] Szent István University, Institute of Environmental Science, Department of Water Management, H-2100 Gödöllő, Páter Károly utca 1. Hungary
[c] Szent István University, Faculty of Mechanical Engineering, H-2100 Gödöllő, Páter Károly utca 1. Hungary
[d] Eötvös Loránd University, Department of Geology, H-1117 Budapest, Pázmány P. stny 1/C., Hungary
[e] Budapest Business School University of Applied Sciences, Faculty of Commerce, Catering and Tourism, Department of Methodology, H-1054 Budapest, Alkotmány utca 9-1, Hungary
[f] Budapest University of Technology and Economics, Department of Sanitary and Environmental Engineering, H-1111 Budapest, Műegyetem rakpart 3, Hungary

ARTICLE INFO

Keywords:
CCDA
Trophic status
Oligotrophization
Sen's trend analysis
Principal component analysis, Lake Balaton

ABSTRACT

The over-enrichment of shallow lakes in nutrients has emerged as one of the main causes of water quality deterioration, and is today a major focus of water quality studies worldwide. In the present work, changes in trophic conditions over three decades (1985–2017) in the largest shallow freshwater lake in Central Europe, Lake Balaton, are assessed using the time series of 10 water quality variables measured at 4 sites, one in each basin of the lake. Using combined cluster and discriminant analyses, and assessing each of the four basins of the lake separately, it was possible to divide the history of the lake into three time intervals. Principal component and Sen's slope analyses highlight the fact that the oligotrophization of the lake took place at a different pace in each of these three major time intervals (1985–1994; 1995–2003; 2004–2017) along the lake's major axis. A significant decrease in the concentration of parameters indicating trophic conditions (e.g. chlorophyll-a, soluble reactive phosphorus) was first observed in the western basins, in the proximity of the main water input to the lake, followed by the eastward spread of this phenomenon. At the same time, the importance of external total phosphorus input to the lake was found to decrease eastwards, thereby diminishing its capacity to explain the variance of the water quality parameters in the lake. Over the time period covered by this study, various measures were taken to reduce the nutrient loads to the lake. These were, in the main, successful, as may be seen in the decade-by-decade overview of the lake's trophic state presented here. A brief review of similar cases from around the world only serves to reinforce the conclusion that a drastic reduction in external phosphorus loads arriving in similar shallow lakes will result in their oligotrophization, albeit with a time-lag of at least ten years.

1. Introduction

Nutrient over-enrichment deriving from intensive anthropogenic activity in the watersheds of lakes has emerged as one of the main causes of deterioration in water quality (e.g. Scheffer, 2013, Schindler, 1974, Schindler et al., 2016, Wetzel, 2001), leading eventually to the degradation of macrophyte vegetation, increased turbidity and, in extreme cases, anoxic conditions (Lau and Lane, 2002). The harmful effects of toxic cyanobacterial blooms endanger aquatic food production and supplies of water for recreation and drinking, leading, in turn, to economic losses, too.

In order to prevent the eutrophication of surface waters, inorganic nutrient inputs must be retained. Evidence shows that, of the inorganic nutrients, it is phosphorus (P) whose retention has the most beneficial effect on the trophic and ecological status of formerly eutrophic lakes (e.g. Sas, 1990; Schindler, 1974; Schindler et al., 2016; Vitousek et al., 2010). Neither can the role of N be neglected, since in estuaries or coastal environments it is a key factor (Carpenter, 2008), and excess reduction of trophic conditions has been achieved by managing not only P but N inputs as well (EPA, 2015). Nevertheless, interventions exclusively aimed at N loads will not lead to the desired oligotrophic states; this can only be achieved by reducing P as well (e.g. Carpenter,

* Corresponding author.
E-mail addresses: hatvaniig@gmail.com, hatvani.istvan@csfk.mta.hu (I.G. Hatvani), kevesolt@geology.elte.hu (J. Kovács), iszekely@geology.elte.hu (I. Székely Kovács), clement@vkkt.bme.hu (A. Clement).

https://doi.org/10.1016/j.ecoleng.2020.105861
Received 7 January 2020; Received in revised form 14 April 2020; Accepted 19 April 2020
0925-8574/

2008; Schindler et al., 2016; Welch, 2009). In spite of the fact that freshwater eutrophication has become a widespread problem over the past half-century and there have been many studies on how to prevent its harmful effects, globally, the number of toxic phytoplankton blooms has continued to increase (Ho et al., 2019; Hudnell, 2008).

A trophic classification of surface waters was first developed in the late 1960s in Europe (Rodhe, 1969), and further developed over succeeding decades. One of the most commonly used indices for the definition of the trophic state of lakes is the trophic state index (Carlson, 1977) relying primarily on the concentration of surface water chlorophyll-a, surface water total phosphorus concentration (TP) and the Secchi depth (Wen et al., 2019). Another widespread classification was formulated in the early '80s by the Organization for Economic Co-operation and Development (OECD), defining the classification of trophic status for freshwater lakes primarily on the basis of the concentration of TP and Chl-a in the water (Vollenweider and Kerekes, 1982). It is these parameters which still constitute the focus of more recently developed models for eutrophication (e.g. Markad et al., 2019, Wen et al., 2019). Therefore, the combined decadal assessment of these parameters is capable of yielding excess information on the effect of external measures aimed at shifting the trophic condition of lakes toward oligotrophization.

One of the most endangered ecosystems in this respect is shallow lakes, which are defined by being well mixed (that is, when subjected to an average wind velocity of 20 km h^{-1} for > 6 h they will mix through their water column (Chapman, 1996)), therefore besides their relatively large surface-to-depth ratio, they are characterized by intense lake-land, air-water and water-sediment interactions (Wetzel, 2001). These interactions render the eutrophication process and formation mechanisms of algal blooms particularly complicated (Qin et al., 2007); they also differ greatly between individual shallow lakes (Janssen et al., 2014). Examples of the adverse effects of algae blooms on shallow lakes have been reported all over the world, e.g. Asia (Qin et al., 2007); North America (López-López et al., 2016; Oberholster et al., 2006); Europe (Hatvani et al., 2014; Sebestyén et al., 2019); South America (Oliveira and Machado, 2013) and Africa (Muli, 1996).

When focusing on the eutrophication of shallow lakes, besides external nutrient loads, the resuspension-desorption of phosphorus from the sediment should also be taken into account, since it plays an important role in the overall nutrient dynamics of shallow lakes (Bloesch, 1995). Indeed, even in the case of reduced external nutrient loads, internal phosphorus load may prevent improvements in lake water quality. At high internal loading, TP concentrations may rise and phosphorus retention can be negative especially in summer (Hatvani et al., 2014; Søndergaard et al., 2003).

Lake Balaton, the largest (surface area 596 km^2) shallow (average water depth 3.2 m) freshwater lake in Central Europe (Fig. 1), has suffered from adverse anthropogenic effects over the last half century (see later and e.g. Hatvani et al., 2014, Padisák and Reynolds, 2003). The lake's watershed is approximately 5180 km^2 (Pomogyi, 1996), and it may be characterized as polymitic. The mean depth and surface area of the lake's geographical basins increases eastwards from 38 km^2 to 228 km^2, while their corresponding sub-watersheds decreases from 2750 km^2 to 249 km^2 (Istvánovics et al., 2007). The largest tributary, the River Zala, which enters the lake at its westernmost and smallest basin, Keszthely Basin (I. in Fig. 1), supplies ~50% of the lake's total water input and accounts for 35–40% of the lake's nutrient input (Istvánovics et al., 2007). The lake's only outflow is the Sió Canal, located at its easternmost end, and this was constructed in the nineteenth century to regulate the water level of the lake.

The accelerated anthropogenic activity (population growth, increasing waste water production, intensified use of fertilizers) in the catchment of Lake Balaton in the second half of the twentieth century resulted in a significant increase in external nutrient load (Hatvani et al., 2015) and a deterioration in the lake's water quality (Sebestyén et al., 2017), and by the end of the 1980s the P load carried by river Zala had doubled (Fig. 2a) compared to the beginning of the 1970s (Herodek et al., 1982; Istvánovics et al., 2007; Sagehashi et al., 2001). For this reason, a regional nutrient load control strategy was worked out for Lake Balaton (Somlyódy and van Straten, 1986), with the most important management measures being: (i) sewage diversion from the eastern and southern shoreline settlements; (ii) the construction of WWTPs in the western part of watershed; (iii) the downsizing of several large livestock farms (Hatvani et al., 2015); and (iv) the construction of the Kis-Balaton Water Protection System (KBWPS) (Hatvani et al., 2011; Kovács et al., 2010; Kovács et al., 2012a), the aim of which was the retention of nutrient loads borought by the Zala River which would have otherwise ended up in Lake Balaton directly; for further details see e.g. Clement et al., 1998; Hatvani, 2014; Hatvani et al., 2014; Hatvani et al., 2015. The combination of these measures and the ten-fold drop in fertilizer usage in the late 1980s (Hatvani et al., 2015) resulted in a TP load reduction of more than 50% compared to the 1980s (Hatvani, 2014). Nevertheless, the oligotrophization of Lake Balaton – and especially its easternmost basin – occurred, albeit with a delay, due to the presence of internal P loads from its sediment (Istvánovics et al., 2004).

As is the case with many temperate shallow lakes, primary production in Lake Balaton was considered to be P limited (Herodek, 1984), while, studies in past decades had focused on the importance of external vs. internal N loads (Présing et al., 2001; Présing et al., 2008). The role of P and N in algal biomass growth was investigated in a way similar to that employed by Schindler (1974) in the Experimental Lakes Area, and it was found that with an increased external P load, algal biomass grew, while N inputs increased the abundance of N-fixing cyanobacteria (Istvánovics et al., 1986). Thus, the more severe limitation of phytoplankton production by P, as compared to that caused by N is also acknowledged in the case of Lake Balaton (Istvánovics and Herodek, 1995; Istvánovics et al., 1986), though it should be recognized that N is found to limit primary production under extreme circumstances, e.g. an abrupt increase in algal biomass (Présing et al., 2008).

With regard to the internal P loads of Lake Balaton, it was found that their maxima are determined by the long-term behavior of the highly calcareous sediment. In the years when the internal P load approaches its maximum, a strong correlation can be observed between the biomass of phytoplankton and the estimated concentration of mobile P, under the influence of the carbonate content of the sediment (Istvánovics, 1988). Otherwise, the biomass of phytoplankton is kept below the highest possible level by physical constraints which depend on hydrometeorological conditions (Hatvani et al., 2014). Consequently, because of this delayed response in lakes (Sas, 1990) it is sometimes hard to find a direct correlation between external load reduction and water quality improvement, particularly over short time periods. In the case of Lake Balaton, thanks to the conscious efforts of the authorities, eutrophication has been successfully managed (Istvánovics et al., 2007). In spite of the global increase in intense lake phytoplankton blooming since the 1980s, in a worldwide study, Lake Balaton remains one of only six lakes, fewer than 10% of the total of 71, which exhibited an internationally acknowledged and statistically significant decrease in trophic status (Ho et al., 2019).

This certainly held true until late summer 2019, when, at the end of August, most of the lake became unexpectedly hypereutrophic due to the bloom of the flagellate *Ceratium furcoides* (Levander) Langhans 1925 and the blue-green *Aphanizomenon flosaquae* Ralfs ex Bornet & Flahault 1886. While both had previously been present in the lake (Padisák, 1985), their blooms had never before produced such an adverse effect. The 2019 bloom first occurred in the Keszthely Basin and spread to the Szigliget Basin, and then to almost the whole lake. The causes are still unclear, but it is obvious that this phenomenon was not to have been expected from the observed trends and the expectations which had formed as a result of the previous success of load reduction measures in and around the lake (Istvánovics et al., 2007). It is suspected that increased temperatures and specific hydrometeorological conditions

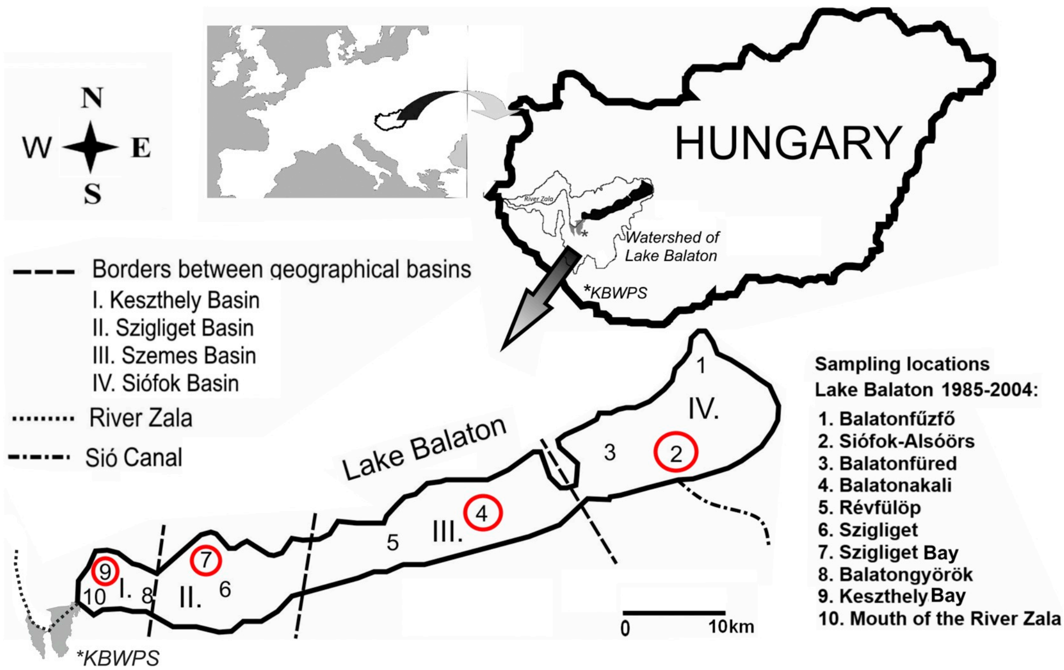


Fig. 1. Lake Balaton, its geographical basins and the 10 sampling sites operated up to 2005 by the responsible water authority. In addition, the Kis-Balaton Water Protection System (KBWPS) and Lake Balaton's watershed is marked on the outline map of Hungary. The sampling sites marked with a red circle (those used in the present work) and site 10 were operating after 2005. The figure is based on that in (Kovács et al., 2012b). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(high water level, etc.) were the cause of this particular bloom (Istvánovics, 2019), while increasing temperatures are set to become an ever-more important factor in eutrophication worldwide (Ho et al., 2019). From a lake management perspective, long term studies of nutrients and their limiting role in primary production are of undoubtedly great importance, since these account for the effects of loading on the natural succession of phytoplankton communities (Istvánovics et al., 1986).

This research aims to offer a long-term overview and an update on changes in the trophic status of Lake Balaton over the last three decades, from the perspective of the multivariate data assessment of its water quality variables. The specific goals are to explore the spatio-temporal changes in the trophic status of Lake Balaton by (i) exploring whether distinct time periods can be distinguished in the history of the lake, (ii) determining the robust trends in parameters indicating trophic status within these time periods in its different basins, and (iii) assessing the change in the importance of external phosphorus loads on general water quality, including trophic indicators.

2. Materials and methods

2.1. Sampling sites and acquired dataset

In the course of the research 10 water quality parameters were selected for analysis. These were measured bi-weekly/monthly at four sampling sites between 1985 and 2017 (that is, the last year for which overall data for the Lake were available), one from each geographical basin of the lake (Fig. 1), a total of approximately 5600 data. The data were acquired from the Central Transdanubian Water and Environmental Inspectorate, and had been collected as a part of the National Water Quality Monitoring System.

Due to the occasionally insufficient number of samples and/or their values being below the level of detection (LOD), the set of parameters analyzed was restricted to the following: soluble reactive phosphorus (SRP), total phosphorous (TP), chlorophyll-a (Chl-a; mg l^{-1}), electrical conductivity (EC; μS cm^{-1}), ammonium – N (NH_4 – N), dissolved oxygen (DO; biological oxygen demand (BOD), chemical oxygen demand (COD), water temperature ($T_{W;}$ °C) and pH. These parameters were chosen so as to provide continuous temporal coverage over the whole of the investigated time interval as they were consistently measured using the same methods and at the same locations. It should be noted that in 2005 the monitoring was spatially recalibrated (Kovács et al., 2012b). In relation to this point, because of these changes, most forms of N had to be omitted from the evaluation. Total nitrogen was not recorded up to 2004, while in 80% of cases, the values of nitrate-N after the recalibration of the monitoring network in 2005 were below LOD. Thus, neither could have been incorporated into the study. However, data concerning ammonium-nitrogen were available for the complete period, and it is this form which is in any case the preferred N form for algal N uptake in Lake Balaton (Présing et al., 2001), as indeed in other lakes, also (Mitamura et al., 1995).

2.2. Methodology

After preprocessing the data (outlier detection and filtering of typos), its descriptive statistics (mean, median, coefficient of variation etc. following Kovács et al., 2012c) were calculated basin by basin to obtain an overview of the dataset. The next step was to use the available indicator variables to shed light on decadal changes in the trophic status of the various basins of the lake following the standard OECD classification (Vollenweider and Kerekes, 1982), which is the most widely accepted worldwide (Istvánovics, 2009). This classification estimates the trophic status of a water body primarily using information on the concentration of the limiting nutrients (TP) and a proxy for phytoplankton biomass (Chl-a). It had been previously used in the case of Lake Balaton (e.g. Crossetti et al., 2013; Istvánovics et al., 2007) and in numerous studies worldwide (e.g. Cloutier and Sanchez, 2007; Marsden, 1989), thereby ensuring the comparability of the results of the present study on a global scale. The question of whether time intervals with common patterns exist in the water quality parameters (WQP) time series was investigated by using combined cluster and discriminant analysis (CCDA Kovács et al., 2014; Section 2.2.1) on the

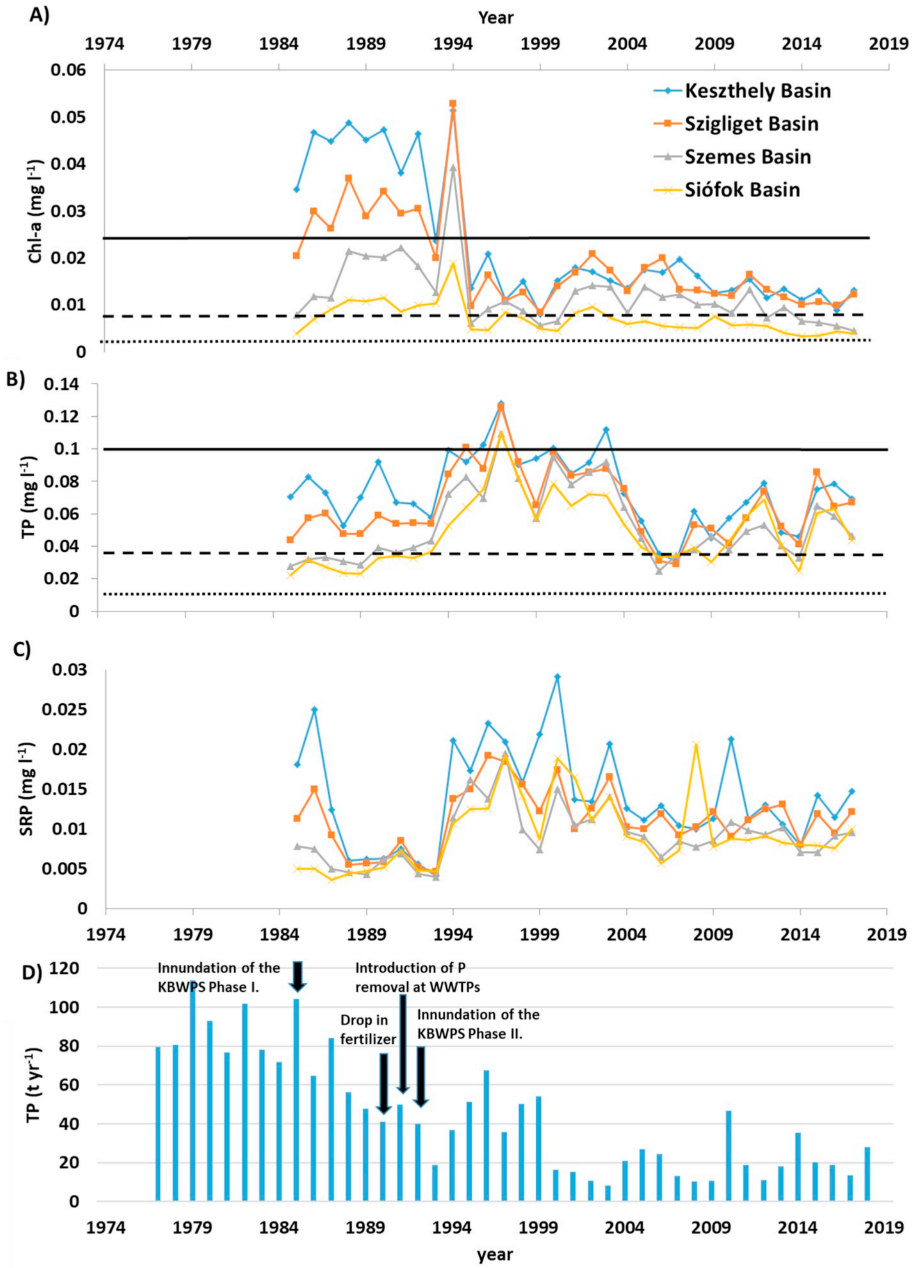


Fig. 2. Annual average concentration of Chl-a A) TP B) and SRP C) for the four investigated basins of Lake Balaton along with the inflow of annual TP loads from the River Zala through the KBWPS to Lake Balaton (sampling location 10 on Fig. 1) D) (redrawn and extended from (Hatvani et al., 2014) C) (1985–2017). In panels A) and B), above the continuous horizontal black line hypertrophic conditions prevail. The dashed line is the threshold for eutrophic conditions, and the dotted for mesotrophic. Below the dotted line oligotrophic conditions prevail, as determined by using the scheme of Vollenweider and Kerekes, (1982).

annual averages of the WQPs. Next, exploratory principal component analysis (PCA; Section 2.2.2) was conducted on the annual averages of the variables in the different basins to explore whether gradually changing common trends might be found as the distance from the lake's main input, the River Zala, increased.

Lastly, the magnitude of change in the concentrations of Chl-a and the P forms was explored for the whole lake basin by basin, using the nonparametric Mann– Kendall test (Kendall, 1975; Mann, 1945) and Sen's slope estimates (Sen, 1968) for the previously obtained differing time intervals using the monthly averaged concentrations of the parameters as the input. In each time interval the annual change in the parameters' values – obtained from the estimated Sen's slope - was given as a percentage of the average concentration of the given parameter in the investigated time interval and basin.

2.2.1. Combined cluster and discriminant analysis

Combined cluster and discriminant analysis (CCDA) is a multivariate data analysis method (Kovács et al., 2014) which aims to find not only similar, but even homogeneous groups in measurement data of known origin (so, in this work, to identify groups of water quality sampling sites). CCDA consists of three main steps: (I) a basic grouping procedure, in this case using hierarchical cluster analysis (HCA), to determine possible groupings; (II) a core cycle in which the goodness of the groupings from Step I and the goodness of random classifications are determined using linear discriminant analysis; and (III) a final evaluation step in which a decision concerning the further iterative investigation of sub-groups is taken. If the ratio of correctly classified cases for a grouping ("ratio") is higher than at least 95% of the ratios for the random classifications ("q95"), i.e. the difference d = ratio–q95 is positive, then at a 5% level of significance, the given classification is not homogeneous in CCDA. Therefore, the division into sub-groups (Step III) and the iterative investigation of these sub-groups for homogeneity is required.

2.2.2. Exploratory principal component analysis

Exploratory principal component analysis (PCA; Tabachnick and Fidell, 2014) was used to find the variables with the greatest influence on the water quality status of the lake over the investigated time period. PCA decomposes the original dependent variables into principal components that explain the original total variance of the dataset component-wise in a monotonically decreasing order. The correlation coefficients between the original parameters and the principal components (PCs) are the factor loadings, and these explain the weights of the PCs in the original parameters, while the PC time series are referred to as PC scores (Olsen et al., 2012; Tabachnick and Fidell, 2014).

In the present case, the input variables for PCA were the annual averages of complete cases (1985–2017) of the WQPs (Sect. 2.1). The Kaiser-Meyer-Olkin (KMO) test (Cerny and Kaiser, 1977) was employed to determine the measure of sampling adequacy (MSA), providing information which allows the decision of whether PCA can be applied to the dataset. The variables with factor loadings outside the ± 0.7 interval were taken as important, while the PCs were taken into account based on their scree plots as suggested by Cattell, (1966) and their eigenvalues, which had to be above 1 according to Kaiser, (1960).

Principal component time series are commonly related to possible explanatory parameters in space (e.g. Magyar et al., 2013; Olsen et al., 2012) and time (e.g. Çamdevýren et al., 2005; Hatvani et al., 2018; Page et al., 2012). In the case of the latter, when correlating the explanatory parameters with the PC scores, the serial correlation of the data should to be considered, since it limits the number of independent observations, not satisfying the assumptions of conventional statistical methods (Macias-Fauria et al., 2012). Thus, in the present study one thousand Monte-Carlo simulations were performed with frequency (Ebisuzaki) domain time series modelling to obtain the correct significance levels of the correlation coefficients (r).

The analysis described in Section 2.2.1 and 2.2.2 was conducted on the annual averages of the complete cases of data considering the lake as a single water body, as well as the four basins of the lake separately. The former approach was justified by the water body designation criteria of the Water Framework Directive (EC, 2000) given that Lake Balaton is one single water body, while latter was justified by the study of Kovács et al. (2012b) highlighting the separate behavior of the four basins of the lake.

R statistical environment was used (R Core Team, 2019) to calculate the descriptive statistics, the Sen's slope estimates with Mann-Kendall tests using the mannKen () function of the wql package (Jassby and Cloern, 2017), and CCDA using the CCDA package (Kovács et al., 2014). The Wilks' λ statistics and PCA were computed using IBM SPSS 26, the statistical significance of the correlation coefficients under serial correlations were calculated using the Windows version of the software provided in Macias-Fauria et al., (2012), and additional tasks were performed in MS Excel 360.

3. Results

To provide an overall picture of the dataset, its descriptive statistics were produced (Table 1). The first quartile and median of the phosphorous forms, Chl-a, BOD and COD were highest in the Keszthely Basin, decreasing continuously with distance from the inlet of the River Zala (Table 1). Conductivity, on the other hand, showed a continuous increase along the main transect of the lake from W to E with respect to medians (from 630 to 666 μS cm^{-1}); this was also true of its first quartile, too (from 680 to 730 μS cm^{-1}). With respect to pH and water temperature, no noteworthy pattern was observed. In general, the highest degree of variability relative to the average in water quality was observed in the Keszthely Basin. However, TP (CV = 85.3) and chl-a (CV = 117.1) were most variable in the Szemes Basin, and orthophosphate in the Siófok Basin.

3.1. Trophic state of Lake Balaton

The relatively high TP and Chl-a values observed throughout the whole investigated time period (1985–2017) made necessary the thorough exploration of the temporal change in the concentration of the indicator variables of trophic status in the lake (Chl-a (Fig. 2a), TP annual mean values (Fig. 2b)) and annual external TP loads arriving in Lake Balaton (Fig. 2d).

It became clear that with regard to the average Chl-a concentration of the whole lake and the Keszthely Basin, 1994 was a turning-point. The average figures for the concentration of Chl-a prior to 1994 were 0.0264 mg l^{-1} and 0.0426 mg l^{-1}, for the lake and the basin, respectively; these figures then dropped by ~60 and ~ 65% in the lake and in the Keszthely Basin, as well as falling below hypereutrophic levels in all the basins (Fig. 2a). Interestingly, in the Szigliget Basin (Fig. 1), phytoplankton biomass in certain years of the early 2000s (e.g. 2002, 2003 and 2005, 2006) and afterwards in 2011 and 2012, was slightly higher and/or comparable to those in the Keszthely Basin. The lowest values for Chl-a were always characteristic of the Siófok Basin (0.008 mg l^{-1}), the furthest from the main external source of TP (Fig. 2d), the mouth of the River Zala.

In almost all cases, TP and SRP values were highest in Keszthely Basin, with the exception of six of the 33 years investigated in the case of the former (Fig. 2b), and four in the latter (Fig. 2c), in which slightly higher values were observed in the neighboring Szigliget Basin. Up to 1994 an increase in TP concentrations characterizes the system, with a parallel decrease in SRP (Fig. 2b,c). Afterwards, TP peaks in 1997 (~0.13 mg l^{-1}) and SRP in, 2000 (~0.03 mg l^{-1}). In the meanwhile, in the Keszthely Basin after 1997 TP concentrations decreased overall by ~50%, to 0.0670 mg l^{-1}, a drop which was even larger in the other basins (Fig. 2b). A classification based on annual averages in the main qualifies the lake's water as eutrophic, and only in the mid-1990s did this become hypereutrophic, as observable from the data of individual basins as well (Fig. 2).

The annual external TP loads arriving in Lake Balaton via the waters of the River Zala for the most part display a continuous decrease, due to the combination of measures taken to reduce nutrient loads in the region (see Sect. 1). However, in the mid-1990s a sudden increase was observed in TP loads reaching Keszthely Bay, while after 2000 the average loads were about 50% of those in the 1990s, with only a couple of years, the first being 2004 (Fig. 2d), with elevated loads due to hydrometeorological conditions (Hatvani et al., 2014). For example, in the region, in 2010 and 2014 the annual precipitation was the sixth and tenth highest in the period 1901 and 2014, amounting to > 875 mm (Jakuschné Kocsis and Anda, 2018; Kocsis et al., 2017).

It should be noted, that the TP loads arriving to Lake Balaton from the River Zala, through the KBWPS (Fig. 2d) resemble the SRP concentrations in Keszthely Basin (Fig. 2c) much more than any other parameter in any of the basins. It was empirically assessed, that after a continuous decrease from 1985, both SRP in Keszthely Basin and TP loads of the River Zala reach a decadal minimum in 1993 (Fig. 2c,d), unlike TP (Fig. 2b). Afterwards, fluctuating in a similar manner, both show peak in 2010, not characteristic of any other basin, or water quality parameter.

Primary indicator variables of trophic status were used to determine thetrophic state of the lake annually (Table A1). Regarding Chl-a, up to 1994 the Keszthely and Szigliget basins were mostly hypertrophic, while the westernmost basins were in most cases eutrophic and in the mid-1980s mesotrophic (Fig. 2a; Table A1) with the exception of 1994, when all the basins were eutrophic/hypertrophic in terms of Chl-a. After 1994 - which would appear to be a tipping point for Chl-a - all the basins except Siófok were mostly eutrophic. In the Siófok Basin, after 1994 a mesotrophic state obtained, which turned exclusive after 2003

Table 1
Descriptive statistics of the water quality parameters in the different basins of Lake Balaton (1985–2017), where M: mean, MED: median, SD: standard deviation, Q1: first quartile, Q3: third quartile; CV: coefficient of variation.

Statistic		Water T (°C)	pH	Cond (μS cm^{-1})	DO (mg l^{-1})	BOD (mg l^{-1})	COD (mg l^{-1})	NH_4-N (mg l^{-1})	TP (mg l^{-1})	SRP (mg l^{-1})	Chl-a (mg l^{-1})
Keszthely Basin	M	16.06	8.49	628.74	10.52	3.01	29.27	0.046	0.079	0.014	0.026
	MED	17.3	8.49	630	10.3	2.55	27	0.03	0.07	0.011	0.016
	SD	7.53	0.21	65.12	2.05	1.64	8.66	0.039	0.046	0.014	0.029
	Q1	9.75	8.37	580	9.16	1.9	23	0.02	0.05	0.007	0.009
	Q3	22.15	8.6	680	11.7	3.9	33	0.05	0.098	0.016	0.031
	CV%	46.9	2.5	10.3	19.5	54.5	29.5	84.5	58.7	95.4	114.8
Szigliget Basin	M	15.54	8.51	636.95	10.43	2.58	25.79	0.046	0.069	0.012	0.021
	MED	16.8	8.5	634	10.2	2.3	24	0.031	0.058	0.009	0.014
	SD	7.39	0.19	63.1	1.82	1.16	6.38	0.037	0.051	0.011	0.023
	Q1	9.4	8.4	585	9.12	1.7	21	0.02	0.04	0.006	0.009
	Q3	21.6	8.62	690	11.7	3.3	28	0.05	0.082	0.013	0.023
	CV%	47.5	2.2	9.9	17.4	45.1	24.7	80.7	73.5	92.7	112.9
Szemes Basin	M	15.36	8.51	644.55	10.35	2.25	22.61	0.041	0.056	0.009	0.013
	MED	16.7	8.52	638.5	10.1	2	22	0.03	0.042	0.007	0.009
	SD	7.36	0.17	63.15	1.89	1.06	4.9	0.033	0.048	0.008	0.016
	Q1	9	8.4	594	9	1.5	19	0.02	0.03	0.005	0.006
	Q3	21.55	8.62	700	11.54	2.8	25	0.05	0.066	0.012	0.015
	CV%	47.9	2.1	9.7	18.2	47.3	21.6	79.5	85.3	86.3	117.1
Siófok Basin	M	14.93	8.52	670.71	10.3	1.99	20.12	0.04	0.05	0.009	0.008
	MED	16.2	8.53	666	10.1	1.8	19	0.03	0.04	0.007	0.006
	SD	7.26	0.16	68.11	1.85	0.94	4.1	0.029	0.038	0.011	0.007
	Q1	8.85	8.4	620	9	1.3	18	0.02	0.027	0.003	0.004
	Q3	21	8.63	730	11.6	2.4	22	0.05	0.061	0.011	0.009
	CV%	48.6	1.9	10.1	17.9	47.3	20.4	72.9	75.3	113.3	92

in terms of Chl-a mean values (Fig. 2a; Table A1).

With regard to TP, the Keszthely and Szigliget basins were mostly eutrophic through the years analyzed, while the Siófok and Szemes basins were in the beginning mesotrophic, later turning eutrophic. The only exception involving all the basins is 1997, in which all of them were hypereutrophic in terms of TP (Fig. 2b; Table 2).

Overall, a clear pattern is visible, with a spatial divide between the western and eastern basins. The western basins (Keszthely and Szigliget) are mostly hypereutrophic at the beginning of the investigated time period (from 1985 to 1994), while the eastern ones (Szemes and Siófok) are eutrophic/mesotrophic. Moving forward in time, this pattern changes to eutrophic for the western basins and dominantly mesotrophic for the eastern basins, with oligotrophic characteristics (in Chl-a maxima) occurring in the mid-2010s in the westernmost, the Siófok Basin (Fig. 2a; Table A1).

3.2. Similarly behaving time intervals and temporal trends of primary trophic indicators

With the use of CCDA on the tagged annual averages of the water quality variables, three similarly behaving ("optimal") time intervals were determined as the basic grouping (Fig. 3). At the $q = 95$ level the biggest difference was 41.2% (Fig. 3), and this split the data into three time intervals: 1985–1994, 1995–2003 and 2004–2017, with the year 1993 not being part of any continuous time interval.

The greatest difference between the basic grouping and a random grouping (Fig. 3: curve d) was observed at the division of the data into three groups (Fig. 3), indicating that this is therefore the optimal grouping. These groups were further divided for the sake of verification, to the point at which all the years become separate, thus demonstrating that homogeneity can only be reached if the separate years form temporal groups alone. However, the three time intervals determined (with similarly behaving years) were objectively determined, and a metric assigned to their existence.

In accordance with the results of the CCDA, it was necessary to investigate the trends for Chl-a and TP as the main WQPs indicating the trophic state of the different basins in each individual basin in turn, and in addition SRP in the three time intervals: 1985–1994, 1995–2003 and 2004–2017 (Table 2). Sen's slopes were determined for Chl-a and P forms, and their annual change relative to the period mean concentration (Mp), representing a significant or insignificant long-term change, was derived (Table 2).

In the first time-period (1985–1994), Chl-a and SRP showed a mostly significant decrease in the western Keszthely and Szigliget basins (Fig. 2a,c), while in the other basins, a significant increase (Fig. 2a,b) and stagnant behavior (Fig. 2c) were observed for Chl-a, TP and SRP, respectively (Table 2). It should be noted here that, although there is a decrease in Chl-a in the Keszthely Basin, the average value is almost 1.5 times higher than in the neighboring Szigliget Basin, and more than twice and almost four times higher than in the eastern basins (Table 2).

The greatest change in all investigated periods and basins was the significant ($p < .01$) decrease of SRP by approx. -7.5% yr^{-1} relative to the period (1985–1994) mean (1.11×10^{-2} mg l^{-1}; Table 2) in Keszthely Basin, resulting in a total 8.2×10^{-3}mg l^{-1} drop in SRP.

Between 1995 and 2003, in the western basins the investigated parameters did not change significantly, although SRP decreased. In the eastern basins, however, Chl-a increased significantly ($p < .05$) by ~4% per year compared to the mean (Mp 1995–2003), while SRP decreased insignificantly (Table 2).

In the last investigated period (2004–2017), SRP concentrations did not show any significant change (Fig. 2c), while Chl-a decreased significantly in all basins (Fig. 2a; between approx. -2 to −4% per year), except Keszthely (Table 2). TP showed a minor (2.5%), but significant ($p < .05$) increase in the Keszthely and Siófok basins (Table 2). The observed trends are all in accordance with the data presented in Fig. 2a-c.

Overall, there is a significant ($p < .01$) decrease in SRP in the western basins in the first of the years investigated, and as this trend weakens over the decades, by the 2010s a decrease in biologically available P and Chl-a presents itself in the eastern basins, although to a degree as yet insignificant (Table 2). It should be noted here that in the Szemes Basin, the −2% annual SRP decrease is significant at $\alpha = 0.1$. It should be further noted that while the trends indicate the change in concentrations within the three distinct water quality time periods, the concentrations of Chl-a should also be considered in the light of the overall change as well (Fig. 2a). This shows a large drop in concentrations, for example, in the Keszthely Basin Chl-a mean

Table 2
Changes in concentration of major trophic indicator parameters in the three time intervals, obtained using CCDA on the different basins. The significance of the Sen's slopes determined by Mann-Kendall tests is indicated by three *** or two asterisks ** for α = 0.01 or 0.05. In any given time period, Mp stands for the mean concentration value. Relative change to the Mp represents the average annual change of the WQP in percentages relative to the Mp of the given time period.

WQP	Keszthely Basin			Szigliget Basin			Szemes Basin			Siófok Basin			Interval
	Chl-a	TP	SRP	Chl-a	TP	SRP	Chl-a	TP	SRP	Chl-a	TP	SRP	
Period mean (Mp)	3.94×10^{-2}	7.38×10^{-2}	1.11×10^{-2}	2.88×10^{-2}	5.65×10^{-2}	8.27×10^{-3}	1.75×10^{-2}	3.86×10^{-2}	6.09×10^{-3}	1.01×10^{-2}	3.1×10^{-2}	5.40×10^{-3}	1985–1994
Relative change to the Mp	−3.8%***	0.0%	−7.4%***	−0.5%	2.1%	−3.2%**	3.2%***	5.9%***	0.0%	4.8%***	4.1%***	0.0%**	
Period mean (Mp)	1.45×10^{-2}	9.95×10^{-2}	1.89×10^{-2}	1.40×10^{-2}	9.28×10^{-2}	1.56×10^{-2}	9.80×10^{-3}	8.43×10^{-2}	1.30×10^{-2}	6.58×10^{-3}	7.55×10^{-2}	1.40×10^{-2}	1995–2003
Relative change to the Mp	0.9%	0.7%	−2.7%	2.4%	−0.8%	−3.0%	4.0%**	1.3%	−2.8%	4.5%**	0.0%	−1.9%	
Period mean (Mp)	1.40×10^{-2}	5.92×10^{-2}	1.22×10^{-2}	1.33×10^{-2}	5.43×10^{-2}	1.01×10^{-2}	9.08×10^{-3}	4.40×10^{-2}	7.72×10^{-3}	4.93×10^{-3}	4.32×10^{-2}	8.05×10^{-3}	2004–2017
Relative change to the Mp	−1.8%	2.6%***	0.0%	−2.4%**	3.2%	0.0%	−4.4%***	1.8%	−2.0%	−3.8%***	2.4%**	−1.3%	

concentrations decreased by ~65% between the period averages of 1985–1994 and 2004–2017 (Table 2).

3.3. Common patterns in the general water quality of Lake Balaton (1985–2017)

Except for the Szigliget Basin, the eigenvalues of the first 3 PCs reached a value of 1, and their cumulative explanatory power fell between ~70 and ~ 80% (Table 3). In the Szigliget Basin, the cumulative variance explained by the first 2 PCs was ~65%. Nevertheless, the explanatory power of the first PC in all the basins was between ~40 and ~ 50%. It is clear that if just the first two PCs are considered for the sake of comparison, their cumulative explanatory power decreases from 68.3% to 57.5% as we move in an easterly direction. According to the KMO test, the MSA yielded values which fell between what may be considered acceptable (> 0.7) and mediocre (> 0.6) in the different basins, thus demonstrating that PCA can provide reliable results.

While the parameters related to biological activity indicating primary production (Chl-a) and its impact on the indices of saprobity (DO, BOD, COD, NH_4-N) had the highest loadings in the first PCs, the phosphorus forms (TP and SRP) were most important in the second. The latter continuously lost its importance as one moves eastwards, with their average loading decreasing continuously from 0.88 to 0.71, while in terms of the loadings of TP in the Szemes and Siófok basins, these fell beneath the 0.7 threshold (Table 3). As for Chl-a, the other main indicator of trophic conditions, its highest loading was observed in the westernmost basin, in the Keszthely Basin (loading of 0.91 in PC1), while in the easternmost basin, at Siófok, it does not even reach the 0.7 threshold in either PCs (Table 3).

The PC scores and the time series of the external TP loads (Fig. 2c) for 1985–2017 arriving to Lake Balaton in the waters of the River Zala (sampling location 10 in Fig. 1) were correlated. It was found that the TP loads – as an external driving parameter – correlated significantly with the first PCs ($r > 0.65$), which were recognized as indicating primary production, rather than with the PCs mainly driven by inorganic nutrients (e.g. P forms; Table 3). The TP loads explain ~60% of the variance of PC1, and this decreases eastwards to ~40% in the Szemes Basin. In the Siófok Basin, the significant linear relationship between the TP loads of the River Zala reaching the lake through the KBWPS and the first PC strengthens. However, in this particular basin the representativity of parameters related to primary production decreased in the first PC (Table 3).

In theory, another viable approach would have been to explore the common patterns/trends in the different basins for the three time periods separately; unfortunately, the matrices serving as the input for PCA were singular, and the KMO test indicated that this particular sub-setting of the dataset would therefore be unsuitable for PCA, with an MSA of < 0.5.

4. Discussion

The spatiotemporal development of the trophic status of Lake Balaton over three recent decades was primarily determined by the complex interplay of its natural internal and external nutrient loads (discussed in Section 4.1), as well as the measures taken to reduce these (Section 4.2). These together make the case of Lake Balaton a unique international example of the result of drastic external load reduction measures in order to ameliorate oligotrophization, and have the potential for application worldwide (Section 4.3).

4.1. Factors behind observed changes and trends

From a spatial perspective, it has been demonstrated that the nutrient content of the water decreases from west to east, resulting in a lower trophic status in the eastern basins (Fig. 2; Table A1). This is due to morphological reasons: the (i) increasing size of the

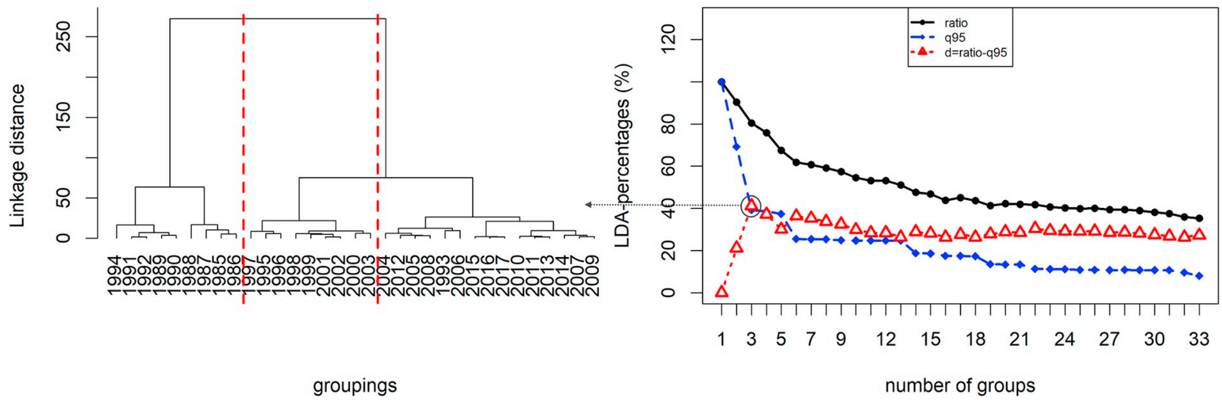


Fig. 3. Dendrogram representing the basic grouping from the initial run of CCDA (left) and the summarized results of CCDA for groupings. Right panel: ratio of correctly classified cases vs. the random classification and the difference values (d).

Table 3
Principal components of the water quality parameters in the different basins of Lake Balaton. Loadings outside the ± 0.7 interval are highlighted in bold. The percentage of original variance explained by the PCs can be found in the penultimate row, and the correlation coefficients of the PC scores and the TP loads of sampling location 10, representing the River Zala arriving to the lake through the KBWPS, in the bottom row. Coefficients of determination significant at $\alpha = 0.01$ or $\alpha = 0.05$ are marked with three ***, or two asterisks ** respectively, taking serial correlation into account.

	Keszthely Basin			Szigliget Basin		Szemes Basin			Siófok Basin		
WQP/PC	PC1	PC2	PC3	PC1	PC2	PC1	PC2	PC3	PC1	PC2	PC3
WT	−0.55	−0.3	0.45	−0.54	−0.51	−0.69	−0.1	0.3	−0.66	−0.36	0.39
pH	−0.3	−0.44	**−0.73**	−0.43	−0.2	−0.38	0.4	0.67	−0.66	0.3	0.34
Cond	**−0.83**	−0.19	0.27	**−0.76**	−0.21	**−0.78**	0.03	0.3	−0.68	0.08	0.26
DO	**0.8**	−0.17	0.04	0.69	0.05	0.43	0.52	−0.08	0.66	0.01	−0.47
BOD	**0.95**	−0.06	0.03	**0.9**	−0.03	0.68	0.47	0.35	**0.76**	0.17	0.28
COD	**0.94**	−0.08	0.03	**0.92**	0.03	0.68	0.49	0.14	0.27	**0.72**	0.37
NH_4-N	**0.76**	−0.03	0.4	**0.81**	−0.06	**0.77**	−0.01	−0.26	**0.8**	0.07	0.08
TP	0.07	**0.85**	−0.25	−0.19	**0.87**	0.57	−0.67	0.4	−0.42	0.66	−0.48
SRP	−0.05	**0.88**	0.08	−0.3	**0.82**	0.53	**−0.75**	0.24	−0.51	**0.71**	−0.25
Chl-a	**0.91**	−0.2	−0.14	**0.82**	−0.13	**0.86**	0.03	0.28	0.61	0.43	0.44
Explained variance	49.30%	19.00%	10.50%	46.50%	18.00%	42.70%	19.10%	11.30%	38.50%	19.00%	12.70%
TP loads (r^2)	0.58**	0.05	0.01	0.48***	0.03	0.42***	0.00	0.18**	0.50***	0.03	0.08

basins, the increasing residence time of the water (Istvanovics, 2002), (ii) decreasing area-specific nutrient loads eastwards (Istvánovics et al., 2007); and (iii) the increasing distance from the mouth of the River Zala, which brings ~50% of the lake's total water and 35–40% of its total nutrient input (Istvánovics et al., 2007) and 90–95% of the nutrient input of the Keszthely Basin (Istvánovics et al., 2004). In the case of the latter, the various interventions intended to reduce its nutrient loads were also a significant factor (see Section 4.2). Taken together, these factors resulted in the decreasing gradient of P and Chl-a content (Tables 1 and A1), influenced to an ever-decreasing extent by the TP input of the River Zala as one heads east on the lake, away from Keszthely, dropping to ~40% (Table 3). Sediment resuspension in Lake Balaton is much higher than in other shallow lakes, moreover, since its carrying capacity is P-determined, its internal loads are of major importance, i.e. the internal P loads can reach the magnitude of external SRP loads in dry years, e.g. early 2000s (Istvánovics et al., 2004), when external loads are smaller than the multidecadal average (Fig. 2d). With the increase in the water's residence time and the previously-discussed eastward gradual decrease in internal and external nutrient loads, the importance of algal biomass in explaining the water quality variability of the lake decreased (Table 3).

From a temporal perspective, although the measures taken (see Section 1) may be considered as the primary factor in the oligotrophization of the lake (Hatvani et al., 2015; Istvánovics et al., 2007), the local hydrometeorological conditions (e.g. precipitation, runoff, temperature, wind) had a measurable effect on its water quality (Hatvani et al., 2014; Istvánovics et al., 2004), as well as, in the form of such phenomena as local temperature, cloud cover, etc., on the water quality of the Kis-Balaton Water Protection System (Hatvani et al., 2017), which serves as a pre-treatment wetland for the loads arriving with the waters of the River Zala (Tátrai et al., 2000). The separation of the three time periods in the history of the lake coincides with major interventions to reduce its loads, as well as larger-scale economic changes (Section 1). However, in certain cases, e.g. the dividing line of 1994/'95, the lake responded to the drop in nutrient loads (Hatvani et al., 2015) with a time-lag; for details, see Section 4.2.

In addition, the reason for the unprecedented behavior of the year 1993 seems likely to be the dry conditions then prevailing (Hatvani et al., 2014), since this was the fifth driest spring of the region in the twentieth century (the spring precipitation amount was only 78.6 mm (Kocsis et al., 2020)). This resulted in decreased external loads, and thus a significant drop in P and Chl-a concentrations throughout the lake (Fig. 2a-c; Table 2). This was also the time of the lowest average SRP concentration (1993 annual avg. 0.0043 mg l^{-1}) in the lake between 1985 and 2017; additionally, this was also the point at which the smallest difference between the basins (max-min SRP in 1993: 0.0007 mg l^{-1}; Fig. 2c) was observable. These conditions were accompanied by a lack of large algae blooms, such as that which occurred in the following year 1994 (Istvánovics et al., 2007).

4.2. The results of the interventions taken to reduce trophic conditions

The significant decreasing trend (Sen's slope on annual averages; $p = 6.83 \times 10^{-5}$) in external TP loads from the watershed of the River Zala (1985–2017; Fig. 2d) was not followed by a direct monotonic decrease in P concentration in the lake as a whole, nor its continuous oligotrophization (Fig. 2b, Table A1, respectively). By way of contrast, SRP mirrored the decrease in external loads to a much greater extent (Fig. 2c,d; Table 2). This may be explained by the following facts:

(i) the lake responds to external load reduction measures (e.g. the inundation of the KBWPS, sewage diversion, P precipitation at WWTPs from the area, see Sect. 1) with a time-lag (Istvanovics, 2002), as in the case of other shallow lakes (Sas, 1990).
(ii) the KBWPS is capable of removing a higher ratio of nutrients from an already elevated input load, while contrariwise, if the loads are reduced, its efficiency also decreases (Clement et al., 1998), leading to a concomitant increase in the importance of internal loads from resuspension-desorption.

The interventions made at the mouth of the River Zala had primarily a local effect, while investments in the development of the waste water infrastructure (Hajnal and Padisák, 2008; Istvanovics, 2002; Istvánovics et al., 2007), and the significant decrease in the use of fertilizers (Sisák, 1993) in the Balaton watershed played an essential role in the decrease of the external nutrient loads in the 30 years covered by this study. The loads dropped by ~75% between 1977 and 1984 and 2004–2018 to ~21 t yr^{-1} in the Keszthely Basin (Fig. 2d), while in the other basins in this figure was already around 50% less as of 2002 (Istvánovics et al., 2007).

The lagged response to the drop in external loads was not only visible in time, but in each basin taken individually, as seen from the Sen's trend estimates (Table 2). While in the Keszthely Basin, the reduction in external loads had an almost immediate effect, in the eastern basins the decreasing trends in P and Chl-a occurred ~10 years later (Table A1). For example, in the Szigliget, Szemes and Siófok basins, only after 2004 did Chl-a start to decrease significantly ($p < .01$), and while reactive phosphorus concentrations stop decreasing in the western basins by the 2000s, at this point these start to show a decreasing trend in the east (Table 2). Findings from the Keszthely Basin indicated a change its trophic conditions from an initial steadily hypertrophic period (1981–1984), to a transient state of hysteresis (1985–1992) (Scheffer, 2013), which concluded in an alternative, less eutrophic state from the mid-1990s (Hatvani, 2014) where most of the eutrophication processes moved upstream to the KBWPS (Hatvani et al., 2014). However, it was from the Keszthely Basin that the unexpected algae bloom of late summer 2019 spread (Istvánovics, 2019), and this demands the further investigation of the specific triggers of this event. The present update on the trophic changes of the lake extends the previous coverage of the thorough analyses on the trophic status of the lake (Istvánovics et al., 2007; Tátrai et al., 2000) by 14 years. It underlines that its trophic state has indeed decreased. Moreover, such updates are crucial, because, at present information on the most recent (2019) algae blooms cannot be immediately provided.

4.3. Global trends in phosphorus load reductions and oligotrophization

Half a century ago, one of the first major reviews dealing with hundreds of studies on all scales was published, and it concluded that the increase in nutrients such as phosphorus and nitrogen originating from external sources are the most likely causes of the eutrophication of lakes (Vollenweider, 1970). At the beginning of the 1970s, lake experiments provided evidence that the reduction of P input is the most effective tool in the reduction of the trophic state and achievement of oligotrophization (Schindler et al., 2016).

There are a number of examples – including that of Lake Balaton – of situations in which efforts to reduce external P loads have resulted in lower TP and Chl-a concentrations and the eventual oligotrophization of a lake's waters (Jeppesen et al., 2005). Recovery has generally been delayed by the internal load, which is in turn dependent on the long-term behavior of sediments (Marsden, 1989; Sas, 1990; Søndergaard et al., 1999). According to these case studies and reviews, in most lakes a new equilibrium was reached after 10–15 years, a period of elapsed time only marginally influenced by the hydraulic retention time of the lakes. With the decrease in TP concentrations, SRP also declined substantially (Jeppesen et al., 2005).

In the case of Lake Balaton, the improvement in water quality was fastest in the Keszthely Basin, which stands in stark contrast to the delayed change in the eastern basins, a difference due to the specific morphometric features of the lake (Istvanovics, 2002). In accordance with the general experience that very large changes in external TP loading were necessary to change the trophic status of a lake (Marsden, 1989), this delayed, but still surprisingly fast recovery was achieved by an external load reduction of around 75% compared to the input when the lake was hypertrophic. As has been observed, "the OECD supports the suggestion that a large reduction in external P loading is necessary to change the trophic status of a lake: a reduction in the annual mean Chl-a concentration across a trophic category requires an approximately 80% reduction in external TP loading" (Vollenweider and Kerekes, 1982). But it is obvious that there is a substantial variation in the load - response relationships of various lakes (Marsden, 1989), and their recovery after nutrient load reduction may be significantly modified by environmental changes such as global warming (Ho et al., 2019), since the effects of global change are likely to run counter to reductions in nutrient loading rather than reinforcing re-oligotrophization (Jeppesen et al., 2005). Also, it is expected that recovery from eutrophication will be more difficult in shallow lakes (Rolighed et al., 2016), and therefore further efforts are needed to arrive at an estimate of the degree of nutrient reduction likely to be required in a future, warmer climate to mitigate eutrophication.

5. Conclusions

The present study provides a 14-year overall update compared to the landmark study of Istvánovics et al. (2007) on the changes and drivers of the trophic status of the largest shallow freshwater lake in Central Europe, Lake Balaton. It highlights the fact that the oligotrophization of the lake took place at a different pace – as indicated by Sen's trend analysis – in the three major time intervals (1985–1994; 1995–2003; 2004–2017) identified in the history of the lake, and what is more, in space along its major axis. At first around the turn of the 1990s the significant decrease in both algal biomass and biologically available phosphorus was observed in the western basins, those in closest proximity to the main water input to the lake, and afterwards spreading east. The stochastic analyses of the linear interrelations of the water quality parameters and the main external P input to the lake, further nuanced this picture. Those showed a continuous decrease in importance of inorganic nutrients (e.g. P forms) driving the general variance of water quality in the lake toward the eastern basins. The overall results indicated that the extent of oligotrophization depended on (i) hydromorphological conditions (ii) the external load reduction measures (e.g. inundation of the lake's pre-reservoir the KBWPS, reduction in fertilizer usage in the watershed, sewage treatment, etc.) of the late 1980s and the 1990s and (iii) local meteorological/basin conditions (e.g. temperature, resuspension of P from the sediment and desorption of SRP).

The findings, in comparison to international case studies highlight the fact that only with the severe reduction of external nutrient loads, and especially in the case of phosphorus, can the oligotrophization of such shallow lakes be achieved. However, due to sediment resuspension, this will occur only with at least a 5–10-year lag in response to the measures taken.

Author contributions

Conceived and designed the study: IGH, JK. Performed the analysis: IGH, VDB, PT, ISzK. Produced the figures: IGH, VDB, PT. Assessed the results: IGH, VDB, AC, JK. Wrote the paper with contributions from all authors: IGH, AC, VDB. Revised the paper with contributions from all co-authors: IGH and AC. We applied the FLAE approach for the sequence of authors; see https://doi.org/10.1371/journal.pbio.0050018.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Authors would like to thank Paul Thatcher for his work on our English versions. The work of IGH was funded by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. JK was supported by the ELTE Institutional Excellence Program (1783-3/2018/FEKUTSRAT), PT was supported by the 'Felsooktatasi Kivalosagi Program' (NKFIH-1159-6/2019), and AC was supported by the the Water sciences & Disaster Prevention research area of BME FIKP-VIZ, all given by the Hungarian Ministry of Human Capacities.

Appendix A. Appendix

Table A1
Annual means and maxima of Chl-a and TP (mg l^{-1}) for the four basins of Lake Balaton (1985–2017) with their corresponding trophic states marked according to the OECD classification (Vollenweider and Kerekes, 1982). Red: hypertrophic; orange: eutrophic; blue: mesotrophic; green: oligotrophic.

Year	Keszthely Basin			Szigliget Basin			Szemes Basin			Siofók Basin		
	Chl-a mean	Chl-a max	TP	Chl-a mean	Chl-a max	TP	Chl-a mean	Chl-a max	TP	Chl-a mean	Chl-a max	TP
1985	0.035	0.099	0.071	0.020	0.052	0.044	0.008	0.018	0.028	0.004	0.009	0.022
1986	0.047	0.202	0.083	0.030	0.120	0.057	0.012	0.034	0.032	0.007	0.013	0.031
1987	0.045	0.085	0.073	0.026	0.064	0.060	0.012	0.020	0.033	0.009	0.014	0.027
1988	0.049	0.160	0.053	0.037	0.137	0.048	0.021	0.064	0.031	0.011	0.025	0.023
1989	0.045	0.160	0.070	0.029	0.085	0.048	0.020	0.064	0.029	0.011	0.020	0.023
1990	0.047	0.147	0.092	0.034	0.098	0.059	0.020	0.079	0.039	0.012	0.032	0.033
1991	0.038	0.099	0.067	0.029	0.074	0.054	0.022	0.062	0.036	0.009	0.016	0.034
1992	0.046	0.262	0.066	0.030	0.174	0.054	0.018	0.083	0.039	0.010	0.030	0.033
1993	0.024	0.043	0.058	0.020	0.039	0.054	0.013	0.021	0.043	0.010	0.016	0.036
1994	0.051	0.203	0.099	0.053	0.193	0.084	0.039	0.137	0.072	0.019	0.087	0.053
1995	0.014	0.041	0.092	0.010	0.024	0.101	0.006	0.015	0.083	0.005	0.013	0.064
1996	0.021	0.055	0.103	0.016	0.061	0.088	0.009	0.032	0.070	0.005	0.010	0.075
1997	0.011	0.036	0.128	0.011	0.027	0.126	0.011	0.025	0.109	0.008	0.017	0.110
1998	0.015	0.030	0.091	0.013	0.032	0.092	0.009	0.017	0.082	0.007	0.016	0.082
1999	0.008	0.021	0.094	0.008	0.014	0.066	0.006	0.012	0.057	0.005	0.010	0.056
2000	0.015	0.066	0.100	0.014	0.046	0.098	0.006	0.011	0.095	0.004	0.009	0.078
2001	0.018	0.056	0.085	0.017	0.044	0.084	0.013	0.039	0.078	0.008	0.020	0.065
2002	0.017	0.050	0.091	0.021	0.078	0.086	0.014	0.049	0.086	0.010	0.025	0.072
2003	0.015	0.053	0.112	0.017	0.063	0.088	0.014	0.044	0.092	0.007	0.017	0.071
2004	0.014	0.037	0.073	0.013	0.024	0.076	0.008	0.020	0.064	0.006	0.018	0.054
2005	0.018	0.048	0.056	0.018	0.040	0.049	0.014	0.038	0.045	0.007	0.015	0.040
2006	0.017	0.052	0.035	0.020	0.053	0.031	0.012	0.027	0.025	0.006	0.009	0.033
2007	0.020	0.057	0.031	0.013	0.024	0.029	0.012	0.034	0.034	0.005	0.008	0.034
2008	0.016	0.039	0.062	0.013	0.033	0.053	0.010	0.022	0.038	0.005	0.010	0.039
2009	0.013	0.028	0.045	0.012	0.023	0.051	0.010	0.036	0.046	0.008	0.021	0.030
2010	0.013	0.039	0.057	0.012	0.025	0.042	0.008	0.015	0.038	0.006	0.017	0.043
2011	0.016	0.041	0.067	0.016	0.044	0.057	0.013	0.040	0.049	0.006	0.009	0.058
2012	0.011	0.026	0.079	0.013	0.026	0.074	0.007	0.019	0.053	0.006	0.010	0.069
2013	0.013	0.026	0.048	0.012	0.021	0.052	0.009	0.018	0.040	0.004	0.008	0.041
2014	0.011	0.021	0.046	0.010	0.022	0.041	0.007	0.018	0.033	0.003	0.006	0.025
2015	0.013	0.029	0.075	0.011	0.036	0.086	0.006	0.016	0.065	0.004	0.006	0.060
2016	0.009	0.017	0.079	0.010	0.027	0.065	0.006	0.008	0.059	0.004	0.006	0.063
2017	0.013	0.041	0.069	0.012	0.022	0.067	0.005	0.011	0.047	0.004	0.009	0.044

References

EC, 2000. Directive 2000/60/European Commission of the European Parliament and of the Council of 23 October 2000 establishing a framework for community action in the field of water policy. Official J. Eur. Comm. L 327, 1–72.

EPA UEPA, 2015. Preventing Eutrophication: Scientific Support for Dual Nutrient Criteria - Information on Preventing Eutrophication: Scientific Support for Dual Nutrient Criteria. EPA 820-S-15-001.

Hatvani, I.G., 2014. Application of state-of-the-art geomathematical methods in water protection: - on the example of the data series of the Kis-Balaton Water Protection System. Ph.D. Thesis. Eötvös Loránd University 110 pp.

Bloesch, J., 1995. Mechanisms, measurement and importance of sediment resuspension in lakes. Mar. Freshw. Res. 46, 295–304.

Çamdevýren, H., Demýr, N., Kanik, A., Keskýn, S., 2005. Use of principal component scores in multiple linear regression models for prediction of Chlorophyll-a in reservoirs. Ecol. Model. 181, 581–589.

Carlson, R.E., 1977. A trophic state index for lakes 1. Limnol. Oceanogr. 22, 361–369.

Carpenter, S.R., 2008. Phosphorus control is critical to mitigating eutrophication. Proc. Natl. Acad. Sci. U. S. A. 105, 11039–11040.

Cattell, R.B., 1966. The scree test for the number of factors. Multivar. Behav. Res. 1, 245–276.

Cerny, B.A., Kaiser, H.F., 1977. A study of a measure of sampling adequacy for factor-analytic correlation matrices. Multivar. Behav. Res. 12, 43–47.

Chapman, D.V., 1996. Water Quality Assessments: A Guide to the Use of Biota, Sediments, and Water in Environmental Monitoring.

Clement, A., Somlyody, L., Koncsos, L., 1998. Modeling the phosphorus retention of the Kis-Balaton upper reservoir. Water Sci. Technol. 37, 113–120.

Cloutier, R.G., Sanchez, M., 2007. Trophic status evaluation for 154 Lakes in Quebec, Canada: monitoring and recommendations. Water Quality Res. J. 42, 252–268.

Crossetti, L.O., Stenger-Kovács, C., Padisák, J., 2013. Coherence of phytoplankton and attached diatom-based ecological status assessment in Lake Balaton. Hydrobiologia 716, 87–101.

Hajnal, É., Padisák, J., 2008. Analysis of long-term ecological status of Lake Balaton based on the ALMOBAL phytoplankton database. In: Nõges, T. (Ed.), European Large Lakes Ecosystem Changes and their Ecological and Socioeconomic Impacts. Springer Netherlands, Dordrecht, pp. 227–237.

Hatvani, I.G., Kovács, J., Kovács, I.S., Jakusch, P., Korponai, J., 2011. Analysis of long-term water quality changes in the Kis-Balaton Water Protection System with time series-, cluster analysis and Wilks' lambda distribution. Ecol. Eng. 37, 629–635.

Hatvani, I.G., Clement, A., Kovács, J., Kovács, I.S., Korponai, J., 2014. Assessing water-quality data: the relationship between the water quality amelioration of Lake Balaton and the construction of its mitigation wetland. J. Great Lakes Res. 40, 115–125.

Hatvani, I.G., Kovács, J., Márkus, L., Clement, A., Hoffmann, R., Korponai, J., 2015. Assessing the relationship of background factors governing the water quality of an agricultural watershed with changes in catchment property (W-Hungary). J. Hydrol. 521, 460–469.

Hatvani, I.G., Clement, A., Korponai, J., Kern, Z., Kovács, J., 2017. Periodic signals of climatic variables and water quality in a river – eutrophic pond – wetland cascade ecosystem tracked by wavelet coherence analysis. Ecol. Indic. 83, 21–31.

Hatvani, I.G., Kirschner, A.K.T., Farnleitner, A.H., Tanos, P., Herzig, A., 2018. Hotspots and main drivers of fecal pollution in Neusiedler See, a large shallow lake in Central Europe. Environ. Sci. Pollut. Res. 25, 28884–28898.

Herodek, S., 1984. The eutrophication of Lake Balaton: Measurements, modeling and management: with 4 figures in the text. Internationale Vereinigung für theoretische und angewandte Limnologie: Verhandlungen 22, 1087–1091.

Herodek, S., Vörös, L., Tóth, F., 1982. The mass and production of phytoplankton and the eutrophication of Lake Balaton. Hidrológiai Közlöny 62, 220–229.

Ho, J.C., Michalak, A.M., Pahlevan, N., 2019. Widespread global increase in intense lake phytoplankton blooms since the 1980s. Nature 574, 667–670.

Hudnell, H.K., 2008. Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs, 619. Springer Science & Business Media.

Istvánovics, V., 1988. Seasonal variation of phosphorus release from the sediments of Shallow Lake Balaton (Hungary). Water Res. 22, 1473–1481.

Istvanovics, V., 2002. P retention as a function of environmental conditions in a large, reed-dominated wetland. Internationale Vereinigung fur Theoretische und Angewandte Limnologie Verhandlungen 27, 4005–4008.

Istvánovics, V., 2009. Eutrophication of lakes and reservoirs. In: Likens, G.E. (Ed.), Encyclopedia of Inland Waters. Academic Press, Oxford, pp. 157–165.

Istvánovics, V, 2019. Discussions on the summer algae blooms of Lake Balaton. In: Public (Hrsg.), . http://www.hirbalaton.hu/. http://www.hirbalaton.hu/.

Istvánovics, V., Herodek, S., 1995. Estimation of net uptake and leakage rates of orthophosphate from 32P-uptake kinetics by a linear force-flow model. Limnol. Oceanogr. 40, 17–32.

Istvánovics, V., Vörös, L., Herodek, S., Tóth, László G., Tátrai, I., 1986. Changes of phosphorus and nitrogen concentration and of phytoplankton in enriched lake enclosures1. Limnol. Oceanogr. 31, 798–811.

Istvánovics, V., Osztoics, A., Honti, M., 2004. Dynamics and ecological significance of daily internal load of phosphorus in shallow Lake Balaton, Hungary. Freshw. Biol. 49, 232–252.

Istvánovics, V., Clement, A., Somlyódy, L., Specziár, A., Tóth, G., Padisák, J., 2007. Updating water quality targets for shallow Lake Balaton (Hungary), recovering from eutrophication. In: Qin, B., Liu, Z., Havens, K. (Eds.), Eutrophication of Shallow Lakes with Special Reference to Lake Taihu, China. Springer Netherlands, Dordrecht, pp. 305–318.

Jakuschné Kocsis, T., Anda, A., 2018. Parametric or non-parametric: analysis of rainfall time series at a Hungarian meteorological station. Időjárás/Quart. J. Hungarian Meteorol. Service 122, 203–216.

Janssen, A.B.G., Teurlincx, S., An, S., Janse, J.H., Paerl, H.W., Mooij, W.M., 2014. Alternative stable states in large shallow lakes? J. Great Lakes Res. 40, 813–826.

Jassby, Alain D., Cloern, James E., 2017. wq: Some tools for exploring water quality monitoring data. R package version 0.4.9.

Jeppesen, E., et al., 2005. Lake responses to reduced nutrient loading – an analysis of contemporary long-term data from 35 case studies. Freshw. Biol. 50, 1747–1771.

Kaiser, H.F., 1960. The application of electronic computers to factor analysis. Educ. Psychol. Meas. 20, 141–151.

Kendall, M.G., 1975. Rank Correlation Methods, 4. Charles Griffin London, UK.

Kocsis, T., Kovacs-Szekely, I., Anda, A., 2017. Comparison of parametric and non-parametric time-series analysis methods on a long-term meteorological data set. Central Eur. Geol. 60, 316–333.

Kocsis, T., Kovács-Székely, I., Anda, A., 2020. Homogeneity tests and non-parametric analyses of tendencies in precipitation time series in Keszthely, Western Hungary. Theor. Appl. Climatol. 139, 849–859.

Kovács, J., Hatvani, I.G., Korponai, J., Kovács, I.S., 2010. Morlet wavelet and autocorrelation analysis of long-term data series of the Kis-Balaton water protection system (KBWPS). Ecol. Eng. 36, 1469–1477.

Kovács, J., Korponai, J., Székely Kovács, I., Hatvani, I.G., 2012a. Introducing sampling frequency estimation using variograms in water research with the example of nutrient loads in the Kis-Balaton Water Protection System (W Hungary). Ecol. Eng. 42, 237–243.

Kovács, J., Nagy, M., Czauner, B., Kovács, I.S., Borsodi, A.K., Hatvani, I.G., 2012b. Delimiting sub-areas in water bodies using multivariate data analysis on the example of Lake Balaton (W Hungary). J. Environ. Manag. 110, 151–158.

Kovács, J., Tanos, P., Korponai, J., Székely, I.K., Gondár, K., Gondár-Sőregi, K., Hatvani, I.G., 2012c. Analysis of water quality data for scientists. In: Voudouris, Kostas, Voutsa, D. (Eds.), Water Quality Monitoring and Assessment. InTech, pp. 65–94.

Kovács, J., Kovács, S., Magyar, N., Tanos, P., Hatvani, I.G., Anda, A., 2014. Classification into homogeneous groups using combined cluster and discriminant analysis. Environ. Model. Softw. 57, 52–59.

Lau, S.S.S., Lane, S.N., 2002. Biological and chemical factors influencing shallow lake eutrophication: a long-term study. Sci. Total Environ. 288, 167–181.

López-López, E., Sedeño-Díaz, J.E., Ruiz-Picos, R.A., 2016. Shallow lakes of the Mexican Central Plateau: assessing their health condition with oxidative stress biomarkers in sentinel organisms. In: Lake Sciences and Climate Change. InTech, pp. 63–79.

Macias-Fauria, M., Grinsted, A., Helama, S., Holopainen, J., 2012. Persistence matters: Estimation of the statistical significance of paleoclimatic reconstruction statistics from autocorrelated time series. Dendrochronologia 30, 179–187.

Magyar, N., Hatvani, I.G., Székely, I.K., Herzig, A., Dinka, M., Kovács, J., 2013. Application of multivariate statistical methods in determining spatial changes in water quality in the Austrian part of Neusiedler See. Ecol. Eng. 55, 82–92.

Mann, H.B., 1945. Nonparametric Tests against Trend. Econometrica 13, 245–259.

Markad, A.T., Landge, A.T., Nayak, B.B., Inamdar, A.B., Mishra, A.K., 2019. Trophic state modeling for shallow freshwater reservoir: a new approach. Environ. Monit. Assess. 191, 586.

Marsden, M.W., 1989. Lake restoration by reducing external phosphorus loading: the influence of sediment phosphorus release. Freshw. Biol. 21, 139–162.

Mitamura, O., Saijo, Y., Hino, K., Barbosa, F.A.R., 1995. The significance of regenerated nitrogen for phytoplankton productivity in the Rio Doce Valley Lakes, Brazil. Arch. Hydrobiol. 134, 179–194.

Muli, J.R., 1996. Environmental problems of Lake Victoria (East Africa): what the international community can do. Lakes & Reservoirs 2, 47–53.

Oberholster, P.J., Botha, A.-M., Cloete, T.E., 2006. Toxic cyanobacterial blooms in a shallow, artificially mixed urban lake in Colorado, USA. Lakes & Reservoirs 11, 111–123.

Oliveira, M., Machado, A.V., 2013. The role of phosphorus on eutrophication: a historical review and future perspectives. Environ. Technol. Rev. 2, 117–127.

Olsen, R.L., Chappell, R.W., Loftis, J.C., 2012. Water quality sample collection, data treatment and results presentation for principal components analysis – literature review and Illinois River watershed case study. Water Res. 46, 3110–3122.

Padisák, J., 1985. Population dynamics of the freshwater dinoflagellate *Ceratium hirundinella* in the largest shallow lake of Central Europe, Lake Balaton, Hungary. Freshwater Biology 15, 43–52. https://doi.org/10.1111/j.1365-2427.1985.tb00695.x.

Padisák, J., Reynolds, C.S., 2003. Shallow lakes: the absolute, the relative, the functional and the pragmatic. Hydrobiologia 506, 1–11.

Page, R.M., Lischeid, G., Epting, J., Huggenberger, P., 2012. Principal component analysis of time series for identifying indicator variables for riverine groundwater extraction management. J. Hydrol. 432–433, 137–144.

Pomogyi, P., 1996. Second Kis-Balaton Conference: Summary of the Research Result on the KBWPS from between 1991–1995. University of Pannonia Georgikon Faculty of Agriculture, Keszthely.

Présing, M., Herodek, S., Preston, T., Vörös, L., 2001. Nitrogen uptake and the importance of internal nitrogen loading in Lake Balaton. Freshw. Biol. 46, 125–139.

Présing, M., Sprőber, P., Kovács, A.W., Vörös, L., Kenesi, G., Preston, T., Takátsy, A., Kóbor, I., 2008. Phytoplankton nitrogen demand and the significance of internal and external nitrogen sources in a large shallow lake (Lake Balaton, Hungary). In: Nõges, T. (Ed.), European Large Lakes Ecosystem Changes and their Ecological and Socioeconomic Impacts. Springer Netherlands, Dordrecht, pp. 87–95.

Qin, B., Liu, Z., Havens, K., 2007. Eutrophication of Shallow Lakes with Special Reference to Lake Taihu, China, 194. Springer Science & Business Media.

Rodhe, W., 1969. Crystallization of eutrophication concepts in northern Europe.

Rolighed, J., Jeppesen, E., Søndergaard, M., Bjerring, R., Janse, J.H., Mooij, W.M., Trolle,

D., 2016. Climate change will make recovery from eutrophication more difficult in shallow danish lake søbygaard. Water 8, 459.
Sagehashi, M., Sakoda, A., Suzuki, M., 2001. A mathematical model of a shallow and Eutrophic Lake (The Keszthely Basin, Lake Balaton) and simulation of restorative manipulations. Water Res. 35, 1675–1686.
Sas, H., 1990. Lake restoration by reduction of nutrient loading: expectations, experiences, extrapolations. SIL Proceedings, 1922-2010 24, 247–251.
Scheffer, M., 2013. Ecology of Shallow Lakes. Springer Netherlands.
Schindler, D.W., 1974. Eutrophication and Recovery in Experimental Lakes: Implications for Lake Management. Science 184, 897.
Schindler, D.W., Carpenter, S.R., Chapra, S.C., Hecky, R.E., Orihel, D.M., 2016. Reducing Phosphorus to curb Lake Eutrophication is a Success. Environ. Sci. Technol. 50, 8923–8929.
Sebestyén, V., Németh, J., Juzsakova, T., Domokos, E., Kovács, Z., Rédey, Á., 2017. Aquatic environmental assessment of Lake Balaton in the light of physical-chemical water parameters. Environ. Sci. Pollut. Res. 24, 25355–25371.
Sebestyén, V., Németh, J., Juzsakova, T., Domokos, E., Rédey, Á., 2019. Lake Balaton: Water Quality of the Largest Shallow Lake in Central Europe, Encyclopedia of Water. John Wiley & Sons, pp. 1–15.
Sen, P.K., 1968. Estimates of the regression coefficient based on Kendall's Tau. J. Am. Stat. Assoc. 63, 1379–1389.
Sisák, I., 1993. Analysis of Diffuse Loads on the Western Watershed of Lake Balaton Using Nutrient Load Estimation Method. Candidate thesis. Pannon Agrártudományi Egyetem, Georgikon Mezőgazdaságtudományi Kar, Keszthely.
Somlyódy, L., van Straten, G., 1986. Background to the Lake Balaton Eutrophication Problem, Modeling and Managing Shallow Lake Eutrophication. Springer, pp. 3–18.
Søndergaard, M., Jensen, J.P., Jeppesen, E., 1999. Internal phosphorus loading in shallow Danish lakes. Hydrobiologia 408, 145–152.
Søndergaard, M., Jensen, J.P., Jeppesen, E., 2003. Role of sediment and internal loading of phosphorus in shallow lakes. Hydrobiologia 506, 135–145.
Tabachnick, B.G., Fidell, L.S., 2014. Using Multivariate Statistics. Pearson Education Limited, Boston, MA, USA.
Tátrai, I., Mátyás, K., Korponai, J., Paulovits, G., Pomogyi, P., 2000. The role of the Kis-Balaton Water Protection System in the control of water quality of Lake Balaton. Ecol. Eng. 16, 73–78.
Vitousek, P.M., Porder, S., Houlton, B.Z., Chadwick, O.A., 2010. Terrestrial phosphorus limitation: mechanisms, implications, and nitrogen–phosphorus interactions. Ecol. Appl. 20, 5–15.
Vollenweider, R.A., 1970. Scientific Fundamentals of the Eutrophication of Lakes and Flowing Waters, with Particular Reference to Nitrogen and Phosphorus as Factors in Eutrophication. OECD Paris.
Vollenweider, R.A., Kerekes, J., 1982. Eutrophication of waters. Monitoring, assessment and control. In: OECD Cooperative programme on monitoring of inland waters (Eutrophication control). Environment Directorate, OECD 154 pp.
Welch, E.B., 2009. Should nitrogen be reduced to manage eutrophication if it is growth limiting? Evidence from Moses Lake. Lake Reservoir Manag. 25, 401–409.
Wen, Z., Song, K., Liu, G., Shang, Y., Fang, C., Du, J., Lyu, L., 2019. Quantifying the trophic status of lakes using total light absorption of optically active components. Environ. Pollut. 245, 684–693.
Wetzel, R.G., 2001. Limnology, Third edition. Academic Press, San Diego 1006 pp.

istvan60_104_23

Journal of Hydrology 617 (2023) 128925

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

ELSEVIER

Research papers

Modeling the spatial distribution of the meteoric water line of modern precipitation across the broader Mediterranean region

István Gábor Hatvani [a,b], Alaa Eddine Smati [c], Dániel Erdélyi [c,d], Gábor Szatmári [e,f], Polona Vreča [g], Zoltán Kern [a,b,*]

[a] *Institute for Geological and Geochemical Research, Research Centre for Astronomy and Earth Sciences, Budaörsi út 45, H-1112 Budapest, Hungary*
[b] *CSFK, MTA Centre of Excellence, Budapest, Konkoly Thege Miklós út 15-17, 1121 Budapest, Hungary*
[c] *Department of Geology, Eötvös Loránd University, Pázmány Péter stny. 1, H-1117 Budapest, Hungary*
[d] *Doctoral School of Environmental Sciences, Eötvös Loránd University, Pázmány Péter stny 1, H-1117 Budapest, Hungary*
[e] *Institute for Soil Sciences, Centre for Agricultural Research, Herman Ottó út 15, H-1022 Budapest, Hungary*
[f] *Department of Physical Geography and Geoinformatics, University of Debrecen, Egyetem tér 1, H-4032 Debrecen, Hungary*
[g] *Department of Environmental Sciences, Jožef Stefan Institute, Jamova Cesta 39, 1000 Ljubljana, Slovenia*

A R T I C L E I N F O

This manuscript was handled by Marco Borga, Editor-in-Chief.

Keywords:
Meteoric water line
Stable isotopes
RMA regression
Machine learning
Precipitation
EMMWL
WMMWL

A B S T R A C T

The covariance of stable hydrogen (δ^2H) and oxygen (δ^{18}O) isotopes in local precipitation (the local meteoric water line – LMWL) serves as a benchmark in isotope hydrological or paleoclimatological applications. However, the isotope hydrometeorological monitoring network is still sparse in many parts of the Mediterranean, and the degree of spatial data coverage is insufficient to address current needs. To remedy this weakness a spatially continuous interpolated geostatistical product of the LMWLs of modern precipitation across the Mediterranean has been developed. The LMWLs of the stations with data for more than four years between 2000 and 2015 were calculated using reduced major axis regression, and then interpolated across the region using statistical and machine learning methods. Random forest interpolation with buffer distance and elevation gave the best performance in the out-of-sample verification, and was thus used to derive the final interpolated product. The slope and intercept of the LMWLs ranged between ~5.9 to 8.2 and −3.9 to 16.1‰, respectively. A detailed comparison with previous local/regional estimations showed that the model presented here concurs with those, albeit with minor deviations in certain regions. With the present results, it then becomes possible to assess how well grounded the 'Eastern- and Western Mediterranean Meteoric Water Lines' (EMMWL and WMMWL) for the 21st century are. In the eastern Mediterranean, the current model shows a slope (~6.9) and the intercept (~15‰) concurring with local studies, but does not reproduce the frequently cited benchmark values of the EMMWL; slope: 8, intercept ≥20‰. The EMMWL may be considered an idealized isotope-hydrological benchmark useful in regional studies; nonetheless, it cannot be taken as a valid representation of the actual empirical description of the δ^2H – δ^{18}O covariance of local precipitation in the eastern Mediterranean. Defining a uniform regional MWL in the western Mediterranean is not supported by the spatial heterogeneity of LMWL parameters based on the present estimations, calling into question the utility of the 'WMMWL' as an isotope hydrological benchmark.

1. Introduction

Water phase changes occur over various spatiotemporal scales in the global hydrological cycle, resulting in relative changes in the spatio-temporal distribution of stable water isotopes (Yoshimura, 2015). The specific isotopic signatures of atmospheric processes such as evaporation, condensation, and deposition in precipitation are then imprinted in terrestrial water bodies, allowing the application of isotopic data to hydrological studies (Gat et al., 2001). The isotopic composition of hydrogen (δ^2H) and oxygen (δ^{18}O) of precipitation (δ_p) is recognized as highly important natural tracer in the water cycle (Bowen et al., 2019; Fórizs, 2003; Gat, 1996; Gat et al., 2001).

Abbreviations: RMA, Reduced major axis; EMMWL, Eastern Mediterranean Meteoric Water Line; WMMWL, Western Mediterranean Meteoric Water Line.

* Corresponding author at: Institute for Geological and Geochemical Research, Research Centre for Astronomy and Earth Sciences, Budaörsi út 45, H-1112 Budapest, Hungary.

E-mail addresses: szatmari@rissac.hu (G. Szatmári), polona.vreca@ijs.si (P. Vreča), kern.zoltan@csfk.org, zoltan.kern@gmail.com (Z. Kern).

https://doi.org/10.1016/j.jhydrol.2022.128925
Received 6 April 2022; Received in revised form 2 December 2022; Accepted 3 December 2022
Available online 15 December 2022
0022-1694/

Describing the linear relationship between δ^2H and δ^{18}O in precipitation is a classical concept of hydrology (Gat, 2005), and is called the meteoric water line (MWL). It is common practice to define a 'Local Meteoric Water Line' (LMWL) for a given location as the best linear fit of all the precipitation data in δ-space (Gat, 2005). A global compilation of the δ^2H and δ^{18}O of fresh water provided a best-fit line, $\delta^2\text{H} = 8 \times \delta^{18}\text{O} + 10$ (Craig, 1961), later called the Global Meteoric Water Line (GMWL). Following an additional ~30-years of data collection, the database of Global Network of Isotopes in Precipitation (GNIP) allowed the GMWL's revision based on long-term amount-weighted annual means as $\delta^2\text{H} = 8.20\ (\pm 0.07) \times \delta^{18}\text{O} + 11.27\ (\pm 0.65)$ (Rozanski et al., 1993).

The GMWL is very useful in hydrology in understanding the origin of modern and ancient ground water and its interactions with surface waters. The GMWL may be considered an amalgamation of many local meteoric water lines (LMWL) which, in turn, vary greatly in their slope and intercepts on regional or local scales (Kendall and Coplen, 2001; Sharp, 2017). These LMWLs are often evaluated in the context of their deviation from the GMWL – e.g. Tappa et al. (2016)-, which represents the "expected" equilibrium relationship, where the slope is assumed to arise from the ratio of equilibrium fractionation factors (Putman et al., 2019).

The GMWL's slope of 8 reflects the average global relationship between δ^2H and δ^{18}O in precipitation, while LMWL slopes documented at values ranging from 4.8 to 10.9 at ~400 sites globally (Putman et al., 2019) may indicate that at least one season is characterized by precipitation affected by nonequilibrium processes. The intercept of the LMWLs has also been documented as varying across the globe from −24‰ to +27‰ (Putman et al., 2019).

The LMWL represents the characteristic distribution of δ_p and can, if determined from long-term observations, serve as a benchmark in isotope hydrological applications such as the estimation of the infiltration period, or the paleoclimatological evaluation of the groundwater isotope signals (Clark and Fritz, 1997). It is common practice, however, to estimate the LMWLs from just short-term investigations – e.g. Bottyán et al. (2017); El Ouali et al. (2022); Kattan (1997); Liotta et al. (2013); Paar et al. (2019) – an approach which is prone to biasing by anomalous climatic events (Longinelli et al., 2006; Vreča et al., 2007; Vreča et al., 2022). If at least 48 months' data for a given station is used, it can attenuate the possible noise caused by such extreme events and interannual variability in estimating the covariance of δ^2H and δ^{18}O (Putman et al., 2019).

At present the Mediterranean region faces an ever-increasing tendency towards water shortage (García-Ruiz et al., 2011; Hoerling et al., 2012; Iglesias et al., 2007; Villani et al., 2022), which will clearly amplify the potential for sociopolitical and economic disruption, e.g., Kelley et al. (2015), and ecological problems (Cook et al., 2016). Isotope hydrological methods have great potential in the improvement of the scientific understanding of the processes governing the regional water cycle, and so, therefore, to assist in knowledge-based water-resource management (Aggarwal et al., 2005). In some parts of the Mediterranean region the isotope hydrometeorological monitoring network was sparse in the past, leading to the initiation of regional campaigns (IAEA, 2005) and further subregional monitoring. The spatial data coverage is, however, still insufficient for many scientific purposes.

Isotope hydrological studies frequently use LMWLs across the Mediterranean, tackling problems in subsurface (Brkić et al., 2020; Christofi et al., 2020; Hssaisoune et al., 2022; Koeniger et al., 2017; Liotta et al., 2013; Trabelsi et al., 2020; Túri et al., 2020; Vasić et al., 2019) or surface hydrology (El Ouali et al., 2022; Gat and Dansgaard, 1972; Serianz et al., 2021). LMWLs taken from relatively remote locations have recently started to grow in numbers, e.g., Boumaiza et al. (2021); Elghawi et al. (2021); Túri (2019), indicating a great demand for such reference data in the isotope hydrological research in the region. Therefore, reliable estimates of the LMWLs are also important at places where direct isotope meteorological monitoring is not in operation. Remote LMWLs can indeed be misleading for local water resource applications, e.g. Serianz et al. (2021).

The LMWLs might be very similar over extended regions, allowing the definition of regional MWLs. Pioneering studies have assessed the precipitation stable isotope composition on subregional scales in the broader Mediterranean realm. Regional meteoric water lines for the eastern- (EMMWL; $\delta^2\text{H} = 8 \times \delta^{18}\text{O} + 22$), and the western parts (WMMWL; $\delta^2\text{H} = 8 \times \delta^{18}\text{O} + 13.7$) were established in the late 20th century and both continue to serve as widely used benchmarks in isotope hydrological studies (Fernández-Chacón et al., 2010; Giustini et al., 2016; Natali et al., 2021; Surić et al., 2018).

Previously determined models can require revision because of the ongoing spatiotemporally varying changes in the hydroclimate, which may be presumed to induce alterations in the isotope-hydrometeorological signal. A further source of difficulty is the existence of various methodological approaches to addressing the question of describing the linear relationship between the water stable isotopes across the Mediterranean (e.g. Argiriou and Lykoudis (2006)), and these may be expected to result in significantly different estimates (Crawford et al., 2014; Marchina et al., 2020; Natali et al., 2021). Moreover, the temporal resolution of the input data and most importantly, the time period the data cover can cause significant differences in the estimates of the slope and intercept of the LMWLs (Putman et al., 2019), as has been specifically documented in certain sub-regions of the Mediterranean; e.g. Krajcar Bronić et al. (2020a); Vreča and Malenšek (2016).

The regional MWL was defined in the western Mediterranean (WMMWL) on the basis of data from ten coastal stations with relatively short-term coverage from the end of the twentieth century (Celle-Jeanton et al., 2001). The origin of the Eastern Mediterranean meteoric water line (EMMWL) is somehow more obscure. In the eastern Mediterranean it was found that the *d* excess, a second-order stable isotopic parameter (defined as $d = \delta^2\text{H} - 8 \times \delta^{18}\text{O}$ (Dansgaard, 1964)) has characteristically elevated values (well above 18‰) (Gat and Carmi, 1970). This observation for Eastern Mediterranean precipitation subsequently received reinforcement: $d >$ 15‰ (El-Asrag, 2005; Rindsberger et al., 1990). To complicate the issue further, decades later, in Gat (1996) this same observation was referred to as "… the eastern Mediterranean–MWL of $d \sim$ 20‰ Gat and Carmi (1970)…".

Regarding the slope of an eastern Mediterranean regional meteoric water line there is no mention of it in Gat and Carmi (1970). A relationship between precipitation δ^2H and δ^{18}O samples has been documented for Israel with slopes of 5 and 6.3 (Gat and Dansgaard (1972): Fig. 3 therein). A potential source of misunderstanding could be the first numeric appearance of the 'rather well known' equation: $\delta^2\text{H} = 8 \times \delta^{18}\text{O} + 22$ in relation with isotope data of surface water in the Upper Jordan Valley (Fig. 9 in Gat and Dansgaard (1972)); this, however, clearly cannot be taken as a figure for precipitation representative of the eastern Mediterranean. Furthermore, Rindsberger et al. (1990) in Sect. 1 states "The δD and δ^{18}O values of the precipitation, in most cases, follow a relationship such as $\delta\text{D} = 8 \times \delta^{18}\text{O} + d$, the so-called … MWL relationship", which is clearly an oversimplification of the matter. Meanwhile, Fig. 7A (δ^2H vs. δ^{18}O plot) in Rindsberger et al. (1990) presents a reference line annotated as "MWL, $d >$ 22‰" without any underlying data or citation, suggesting that this is just an arbitrarily chosen benchmark in that particular study. At the turn of the twenty-first century, it was reported that rain events of > 20 mm at a monitoring station in Israel between 1995 and 1997, followed a "trend with a slope of ~8, and d-excess of 20 to 30‰" (Ayalon et al., 1998). This, however, cannot be considered an MWL, since as noted above, the LMWL is defined as the best linear fit of all the precipitation data in δ-space (Gat, 2005).

It seems that some studies have taken it for granted that the slope is 8 and the intercept is 22‰ of the EMMWL, although these parameters were not obtained obeying the rules of how meteoric water lines should be derived; see Gat (2005). What is more, further confusion regarding the Mediterranean regional meteoric water lines may have originated from the occasional usage of ad-hoc terminology and acronyms. For example, the Mediterranean Meteoric Water Line (MMWL (Ayalon et al., 1998; Koeniger et al., 2017)), which is used to represent location(s) in the eastern Mediterranean, or the abbreviation EMWL (e.g. Bajjali and Abu-Jaber (2001); Yüce (2005)). Such usage further increases the possibility of confusion over regional MWLs. In the present work, the abbreviations EMMWL and WMMWL are used exclusively to refer to the regional definitions.

A recent study has addressed the geographic variations in the slope of LMWLs across Europe focusing on the temperate and boreal regions of the continent (Lécuyer et al., 2020); it does not, however, cover the Mediterranean region. Thus, the current work aims to provide a spatial extension of that analysis with respect to the slope of LWMLs. Additionally, the present study represents a methodological extension, because the intercept of the LMWL is also evaluated. Moreover, in doing so, the performance of various interpolation methods (classical statistical and machine learning) with different set of predictors (altogether 14 approaches) is investigated in modeling the spatial distribution of the slope and intercept of the local meteoric water lines across the Mediterranean taking into consideration the climatic and geographic context of the region.

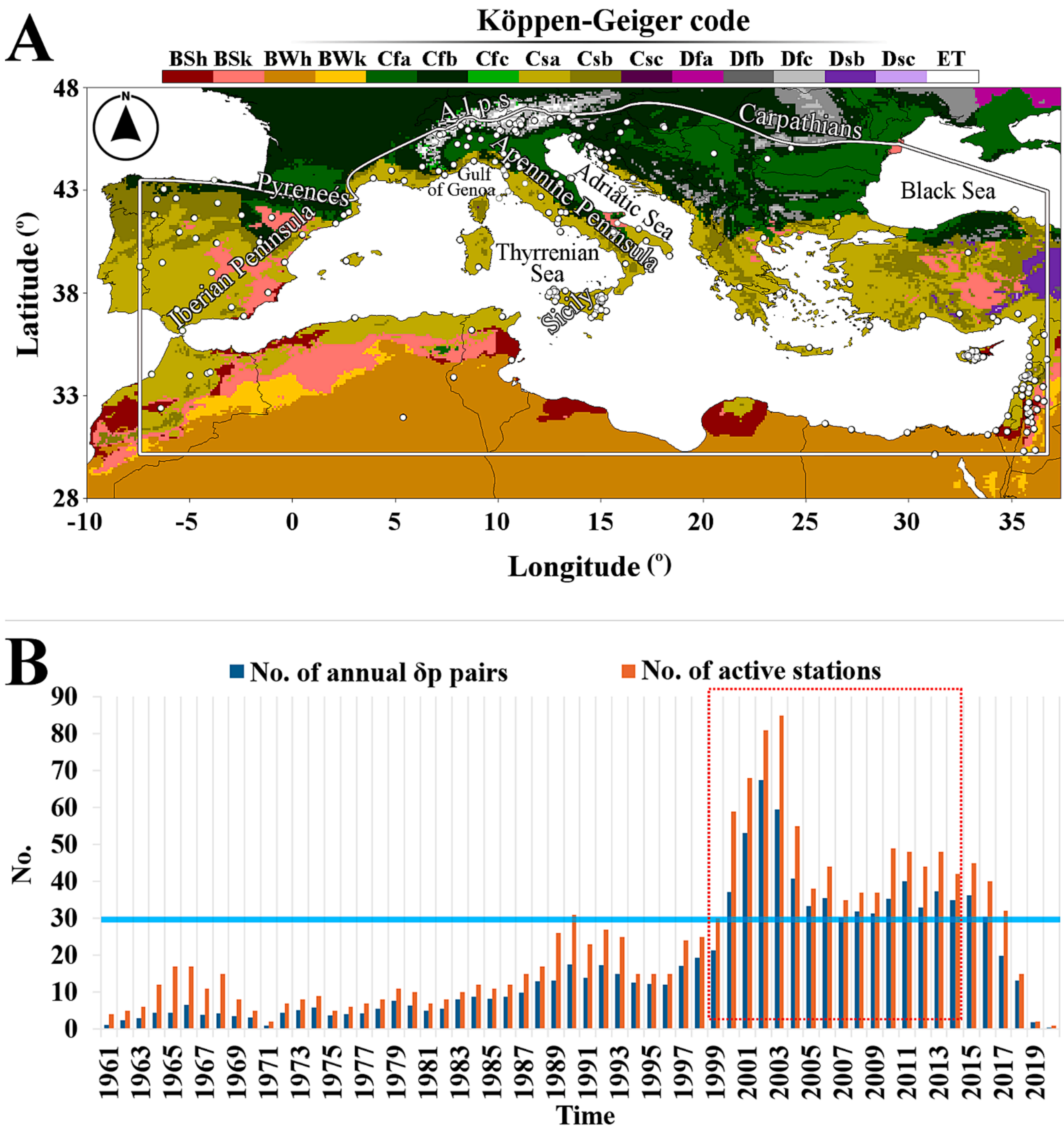


Fig. 1. Study area, monitoring sites and available data. Map of Köppen climate zones (Kottek et al., 2006) and precipitation monitoring sites (white dots) of the region studied (A). Number of δ_p records obtained from the GNIP stations and other data sources (see Sect. 2.2) for the period 1960–2020. The red dotted rectangle represents the focus period for data preprocessing (2000–2015; see Sect. 2.3.1). The horizontal blue line indicates the threshold chosen to select the focus period (B).

2. Materials and methods

2.1. Study area description

The study area covers the Mediterranean Basin from north Africa, across southern Europe to the Middle East. This area extends from 7.5° W to 36.7° E, and from 29.9° N to the orogenic belt from the Pyrenees through the Southern Alps to the Carpathian chains (Fig. 1A).

The Mediterranean climate of the region is characterized by temperate dry hot summers and humid winters, the Csa climate subtype in the Köppen-Geiger climate classification system (Kottek et al., 2006). Warm dry summers (Csb) are more frequent on the coasts close to the Atlantic Ocean in the West and the Black Sea in the East. The southern part of the basin is dominated by arid hot and cold steppe conditions (BSh, BSk) and arid desert conditions (BWh), especially south of 31° N under the influence of the Sahara. North of 41° N, a temperate climate dominates, without a dry season, and with warm (Cfb) and hot summers (Cfa), especially around the Black and Adriatic Seas.

The Mediterranean Basin is characterized by specific patterns of atmospheric circulation, with dry and cold continental air masses that interact with a marine basin characterized by high evaporation rates and relatively high sea surface temperatures (≈ 20 °C) (IAEA, 2005). Such air-sea interactions are important sites of cyclogenesis in the Mediterranean (Bartholy et al., 2009; Kelemen et al., 2015). Compared to the rest of the European continent, which is dominated by Atlantic moisture, the Mediterranean has a more complex moisture source structure, the primary source region being the Mediterranean Sea itself, at least seasonally (Ciric et al., 2018). Previous subregional isotope hydrometeorological assessments have revealed that the variable physiographic and hydrometeorological factors act differently and result in a remarkable divergence in values of δ_p between the different parts of the Mediterranean Basin (Fischer and Mattey, 2012; Giustini et al., 2016; Hatvani et al., 2020; Hunjak et al., 2013). Consequently, the δ^2H – δ^{18}O relationship has also been found to display considerable subregional differences; e.g. Dotsika et al. (2010); Giustini et al. (2016); Vreča et al. (2006).

2.2. Data used

Monthly δ_p values were acquired from a total of 249 monitoring stations (covering 1960–2020), primarily from the Global Network of Isotopes in Precipitation (GNIP), with six stations updated with more recent data: Râmnicu Vâlcea (Varlam et al., 2021), Zagreb (Krajcar Bronić et al., 2020a), Plitvice (Krajcar Bronić et al., 2020b), Ljubljana and Portorož (Vreča et al., 2015; Vreča et al., 2011; Vreča et al., 2008; Vreča et al., 2014; Vreča et al., 2022), and Pisa (Natali et al., 2021) (Table S1). In addition, other local and regional networks were included:

- 11 stations in Sicily operating between February 2002 and March 2003 (Liotta et al., 2006); 6 stations (February 1992 to December 1996) (Longinelli and Selmo, 2003); 8 stations (January 2002 – December 2004) scattered across northern Italy (Longinelli et al., 2006), 3 stations from the Mt. Vulture region (Paternoster et al., 2008), and 7 stations from Tuscany (Natali et al., 2021). In addition, single records from Forni di Sopra (2005–2010) (Cervi et al., 2017) and Riva del Garda from February 2007 to January 2008 (Longinelli et al., 2008);
- data collected at 6 stations in Lebanon between October 2012 and May 2013 (Koeniger and Margane, 2014);
- 6 stations in Hungary operating between January 2005 and November 2017 (Fórizs et al., 2020);
- 3 stations in Croatia operating between May 2012 and September 2013 (Paar et al., 2019) and one station between 2017 and 2019 (Marković et al., 2020);
- Dumbrava (Romania) from April 2012 to July 2015 (Bojar et al., 2017);
- 16 stations in Cyprus (2014–2017) (Christofi et al., 2020);
- 1 station in Serbia (1992–1997 and 2003–2005) (Golobočanin et al., 2007); and
- 2 short time series from Morocco from June 2010 to November 2011 (Wassenburg et al., 2016).

In general, the data gathered were unequally distributed in space (Fig. 1A) and time (Fig. 1B). Until the mid-1980s, the number of active precipitation monitoring stations was limited ($n < 12$). From the mid-1980s onwards their number rose progressively, reaching about 50 by the turn of the millennium. In the early 2000s, owing to a coordinated research project of the IAEA (IAEA, 2005), the network was quite extensive. For instance, in 2002/3 the number of active stations peaked at over 80, while the average number of annual δ_p pairs was ~60 at that time (Fig. 1B). In the 2000s the same number of δ^2H and δ^{18}O data were generally available, although in certain cases the δ^{18}O records were more abundant.

To maximize the spatiotemporal coverage of the δ_p values available from the study region, a focus period of 2000–2015 was chosen. This period was characterized by an annual average number of stations > 31 providing data in any given year. The only exception was 2007, in which the average number of active stations was only 30.2 (Fig. 1B). Nevertheless, the focus period chosen (2000–2015) was not cut in half on account of this one year.

2.3. Methodology

2.3.1. Data preprocessing

As a first step, local Moran's I statistics (Moran, 1948) were calculated to investigate the correlations within different spatial units and their surrounding spatial lags (up to 501 km). This was done for each month from January 2000 to December 2015 for both δ_p, using Moran scatterplots (Anselin, 1996), classifying the spatial autocorrelation (Moran Is) of the stations into four types, and enabling an advanced screening for outliers.

If the Moran scatterplot suggested a possible error, the d-excess values ($d-excess = \delta^2H - 8 \times \delta^{18}O$ (Dansgaard, 1964)) were also compared between the neighboring. If the d-excess values of the possible outlier deviated inexplicably from its neighbor(s) (e.g. diff. > 10), the δ_p values were dropped; cf. Erdélyi (2023). This approach is more complex than that employed in previous studies, which used a static d-excess cutoff threshold (Bowen, 2008; Nelson et al., 2021), outside of which δ_p values were excluded without any consideration of the spatial relationship.

As a result of this approach, δ_p values from Patras (Dec 2000), Ciudad Real (Feb 2002), Antalya (Mar 2003) and Murcia (Jan 2004) were dropped, in all cases the d-excess was < -20. In addition, curiously in 2010 the exact same δ_p values were recorded at the Ankara and Antalya stations and it was not possible to trace back the origin of the GNIP database error or decide to which station the values belong. Thus, these were excluded from the analyses.

Next, the monitoring stations were classified into three quality categories (QC) according to the number of data recorded at each one during 2000–2015.

- Stations belonging to QC 1 had to have δ^2H–δ^{18}O pairs from > 48 months, as this was the amount of data required for any LMWL calculation for it to take on the characteristics of the long-term distribution without bias caused by natural noise (Putman et al., 2019).
- In the study area, in high summer, precipitation may not fall for months. Thus, an additional 'quality category' was set up. QC 2 had to have data from at least four years covering not < 16 months, because below this threshold the use of RMA regression frequently resulted in insignificant relationships (see Sect. 2.3.2).
- The stations not meeting any of the criteria above (QC 3) were excluded from the analysis.

The possibility was entertained that stations situated close to each other (within 10 km) could be merged to improve the input data for LMWL estimation. Since certain monitoring sites had been relocated within a short distance, or restarted at a nearby locality after a multi-annual halt in monitoring, the datasets for such nearby stations were merged (Table S1). The coordinates of the one functioning longest from among those merged was used in the final dataset. In all cases the inter-station distance of the stations forming these merged datasets was smaller than 8 km. The final quality-controlled dataset contained 42 QC1 and 20 QC2 stations, of which 11 had been merged.

2.3.2. Reduced major axis (RMA) regression

Traditionally the LMWLs are defined employing ordinary least squares regression (OLS) or reduced major axis (RMA) regressions (IAEA, 1992). The conceptual difference between the two approaches is that while OLS regression minimizes the sum of the squared deviations of the predicted values, RMA minimizes both the vertical and the horizontal distances between each point and the regression line, i.e. it uses diagonal residuals that have slopes opposite to the slope of the regression line (Clarke, 1980). Since RMA is less prone to outliers, and more importantly, the relationship of the two assessed variables can be described by physical laws (Carroll and Ruppert, 1996), the application of RMA is more appropriate to the definition of the LMWL (Crawford et al., 2014; Putman et al., 2019).

A more advanced option could have been to use precipitation weighted RMA (PWRMA) regression (Crawford et al., 2014), but that would require precipitation amounts recorded for all δ_p values, a requirement not met in the present case. Moreover, studies support that the difference between slope and intercept values estimated by RMA and PWRMA agree within error bounds (Boschetti et al., 2019; Krajcar Bronić et al., 2020a). Therefore, in the present study RMA regression was applied to raw monthly δ_p values using the `lmodel2` package (Legendre, 2018) in `R` (R Core Team, 2019) with 2,000 permutations to obtain the significance values.

2.3.3. Interpolation techniques

Three major types of interpolation techniques were employed in the study, with different sets of predictors to obtain a spatially continuous map of the intercept and slope of the RMA regression. Inverse distance weighting (IDW), a traditional interpolation method (Webster and Oliver, 2008), and two machine learning (ML) methods, specifically random forest (RF) (Li et al., 2011; Liu et al., 2012) and support vector machine (SVM) (Cortes and Vapnik, 1995), were used.

IDW interpolation is a method in which each point of estimation is calculated as the average of nearby sample values, weighted by their distance from the estimation point. It does not make assumptions about spatial relationships, except for the basic assumption that nearby points are expected to be more closely related to the value at the estimation location than distant points. Two alternatives were trialed, linear (weighting power $\alpha = 1$; IDW_{p1}) and nonlinear (weighting power $\alpha = 2$; IDW_{p2}).

RF is a machine learning algorithm. In the procedure, a combination of a series of tree structure predictors is assessed. Here, each tree depends on the values of a random vector sampled independently with the same distribution for all trees in the forest. The generalization error for forests converges to a limit as the number of trees in the forest increases. The generalization error of a forest of tree classifiers depends on the strength of the individual trees in the forest and the correlation between them (Breiman, 2001). Its predictions are based on the average results of the decision trees, which use bootstrap sample (bagging) to eliminate the possibility of over-fitting; for details see e.g. Biau and Scornet (2016); Breiman (2001); Prasad et al. (2006). As with RF, SVM is another frequently used supervised machine learning technique for classification and regression purposes. However, the theory behind SVM is different to that of RF. It relies on the application of Kernel functions, which also means that SVM is a nonparametric technique, in finding a function that deviates from an observation by a value not greater than a given threshold value for each observation and at the same time is as flat as possible. As a consequence, the function fitted by SVM is the hyperplane that has the maximum number of observations (Cortes and Vapnik, 1995).

Six combinations of predictors which can influence the stable isotopic composition of precipitation, consequently the slope and intercept of the LWMLs in the region, were considered in the case of the ML algorithms:

- latitude, longitude (in Web Mercator EPSG: 3857), Köppen-Geiger climate class codes (category variable) with the approaches denoted as RF_{KG} & SVM_{KG},
- latitude, longitude and elevation (ele), denoted as RF_{ele} & SVM_{ele}, elevation was incorporated since the local orographic features may significantly change the isotopic composition of precipitation in the Mediterranean (Liotta et al., 2006; Liotta et al., 2013),
- latitude, longitude, Köppen-Geiger climate class codes, and elevation, denoted as $RF_{KG\text{-}ele}$ & $SVM_{KG\text{-}ele}$,
- the distance matrix of the monitoring stations, and again the Köppen-Geiger codes following the RFsp procedure of Hengl et al. (2004), with the approaches denoted as $RF_{sp\text{-}KG}$ and $SVM_{sp\text{-}KG}$,
- the distance matrix of the monitoring stations, and elevation, denoted as $RF_{sp\text{-}ele}$ and $SVM_{sp\text{-}ele}$,
- the distance matrix of the monitoring stations, the Köppen-Geiger climate class codes and elevation, denoted as $RF_{sp\text{-}KG\text{-}ele}$ and $SVM_{sp\text{-}KG\text{-}ele}$.

The IDW interpolation was performed using Golden Software Surfer 15 and GS+ 10. In the RF approach, the value of the `nodesize` and `mtry` parameters was selected automatically using the `tune` command of the `randomForestSRC` package (Ishwaran et al., 2021; Ishwaran et al., 2008) in `R` (R Core Team, 2019), while for RF_{sp}, these were chosen as 5 and 10 for the slope and intercept parameters, respectively, using the `ranger` package (Wright and Ziegler, 2015). The SVM modelling was carried out using the `caret` (Kuhn et al., 2020) and `e1071` (Meyer et al., 2019) `R` packages.

2.3.4. Validation

The results of the different interpolation schemes were validated by forming two pairs of training and validation datasets. Validation stations were only selected from those countries where there was more than one station available (not considering merged stations). The Validation-set 1 (Vs1; $n = 10$) consisted of one randomly chosen station from each country, preferably from its Q2 station(s). Validation-set 2 (Vs2; $n = 10$) consisted of the first stations in alphabetical order from each country (Table S1). The reason the stations were chosen from different countries was to avoid biased validation results that might arise due to some spatial clustering.

The sets of stations used in the actual interpolation were complementary to Vs1 and Vs2, and were referred to as Set 1 and Set 2, respectively. After applying the interpolation methods to estimate the slope and intercept of Set 1 and Set 2, the squared point-differences for stations in Vs1 and Vs2 were calculated. The performance of the 14 interpolation schemes was primarily evaluated using the squared errors of Vs1 and Vs2, while visual inspection of the interpolated maps was employed to check for systematic interpolation errors, artefacts etc. First and foremost, the median of the squared errors was taken into consideration rather than the mean, since latter is much more sensitive to outliers (Wilcox, 2003), which is a crucial strength in the case of a relatively small sample size, as with the present one.

Journal of Hydrology 617 (2023) 128925

3. Results and discussion

3.1. Local meteoric water lines in modern precipitation across the Mediterranean

In the estimations presented here, the point values of the stations and the pattern seen on the interpolated maps are compared with previous studies of the region. The slope of the δ^2H – δ^{18}O relationship varied between ~5.9 and 8.3. The lowest values, some even falling below 7, were observed in the east and in the western parts and N Africa. The highest values, exceeding 8, in the north, in the Alpine foothills of Italy (Fig. 2A).

Considering subtropical arid or seasonally hot and dry regions (Köppen-Geiger climate class B and Cs), ~40% of the slopes obtained fall within the typical range (5 to 7) reported for these climate regions in a global assessment (Putman et al., 2019). The rest of the stations assigned to these climate classes had slightly steeper slopes (max Lake Massaciuccioli: 7.9 in zone Csa), which is still not unusual in these climates. Considering humid temperate and seasonally snow dominated regions (Köppen-Geiger climate class Cf and D), 77% of the obtained slopes fall within the typical range (7 to 8) for these climate regions reported in a global assessment (Putman et al., 2019). The remainder is in the vicinity of the southern Alps, where, as previously mentioned, the highest slope values are observed. In the subtropical arid or seasonally hot and dry regions (Köppen-Geiger climate classes B and Cs), almost all (90%) of the intercept values fall within the typical range (0 to 15‰) reported for these climate regions in (Putman et al., 2019). The outliers comprise mainly stations with low intercepts in central Iberia. In the case of humid temperate and seasonally snow dominated regions (Köppen-Geiger climate classes Cf and D), only 14% of the intercept values obtained fall within a typical range (−20 to 5‰) (Putman et al., 2019). This discrepancy may be due to the differing extent of various colder climate zones incorporated into the studies. It should be noted that the stations studied here are characteristically rather humid, and experience less snowfall (Fig. 1 and Table S1), whereas snowfall tends to result in negative intercepts (Putman et al., 2019).

The large-scale spatial patterns fit well into the estimations made for a continental transect across Europe, where slopes ranged from 6 to 10 (Lécuyer et al., 2020). In fact, from the western margin of the area studied up to the Adriatic (~12°E), a similar longitudinal trend is observed as that reported across the temperate and boreal regions of Europe, reaching a slope of 8 Lécuyer et al. (2020): Fig. 4 therein. Further eastward, however, while the slope of the relationship assessed from the continental transect continues to increase up to 9 (Lécuyer et al., 2020), in the case of the Mediterranean, it declines.

The spatial pattern of the intercept, on the other hand, shows a more complex pattern. In central Iberia, the intercept is $<$ 3.5‰, while the coastal parts have values ranging between 2.7 and 6‰. Then, after a gradual increase, higher values are seen surrounding the Adriatic Sea, peaking at $>$ 10‰ in northeastern Italy and Slovenia (Fig. 2B), matching the area of the steepest slopes (Fig. 2A). Eastward from the Adriatic, a gradual decrease is observed. Around 17°E intercepts of 7.5‰ are observed to decrease to around 5‰ (20–23°E). A subregional pattern is seen surrounding the Black Sea, where intercept values are $<$ 3.2‰ (Fig. 1A). It is a common climatic feature of the 'low-intercept' areas (both central Iberia and this particular region) that the summer is characterized by very dry conditions (Table S1). An explicit contrast is observed in the southeastern sector of the Mediterranean Basin, where

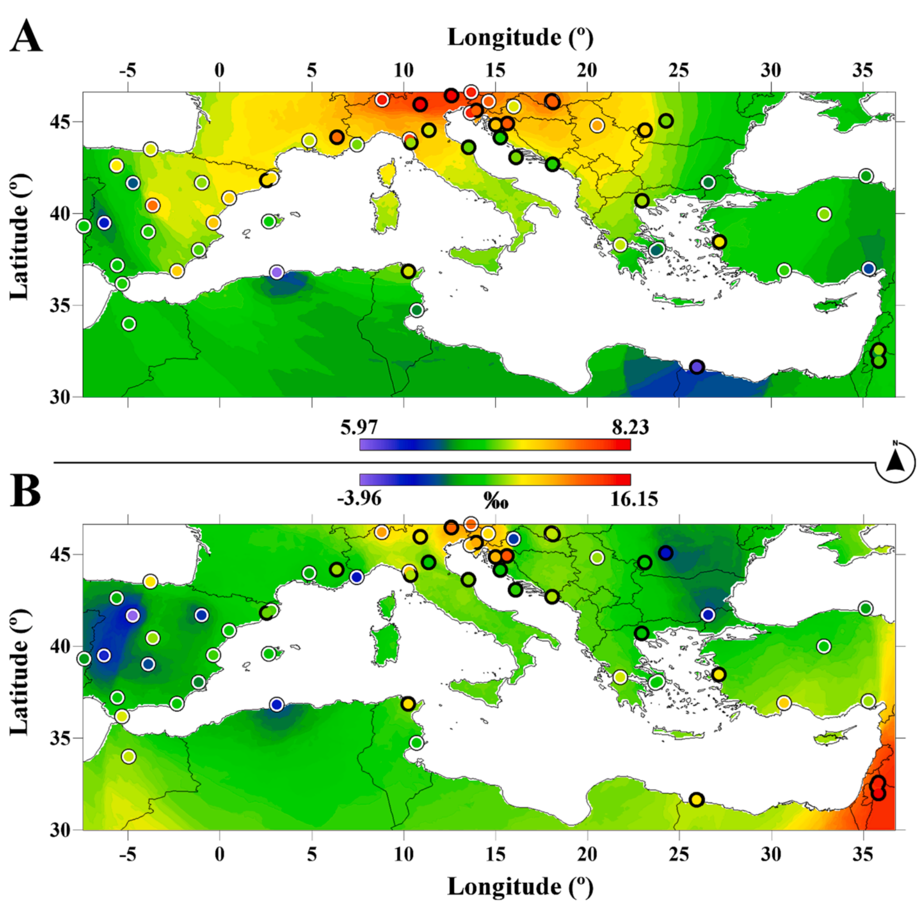


Fig. 2. Slope (A) and intercept (B) of the $RF_{sp\text{-}ele}$ interpolated map of the Mediterranean LWMLs. The colored circles with a white outline indicate the LMWLs of the precipitation monitoring stations belonging to QC1, while tha black outline denotes QC2 stations. The GeoTIFF versions of the maps can be found as Figs. S1 and S2 for the slope and the intercept, respectively.

altogether the highest (> 16‰) intercept values are to be found (Fig. 2B).

In the northern part of Africa, the number of stations used to estimate the LMWLs and interpolate those is smaller and much less continuous compared to the other regions in the Mediterranean (Fig. 2). Therefore, attention should be paid to these points when the estimations for this specific sub-region are used.

3.2. Performance evaluation of the spatially continuous estimations of the LMWLs across the Mediterranean region

In choosing the best performing interpolation method, the median, the upper quartile and the interquartile interval obtained squared differences and the question of how prone the method was to showing extreme outliers were all considered. If the number of squared differences classified as extreme outliers (Fig. 3) is ≥10% (practically $n = 2$) for either the slope or the intercept, said method is discarded. This was true for all the SVM approaches regardless of the predictors applied. Since the SVM methods gave the most extreme outliers in the case of the slope and intercept, those were not considered in the following. It was also determined that there was no significant difference either between the medians as calculated using Mood's independent samples median test (Mood, 1950) (asymptotic significance $p = .549$ and .72 for the slope and intercept, respectively), or in the distribution of the squared errors between the three models with the best performance based on independent samples using the Kruskal-Wallis test (Kruskal and Wallis, 1952), which yielded values of $H(2) = .663, p = .718$ for the slope and $H(2) = .82, p = .644$.

The inverse distance approaches performed worst in the case of the slope regarding the median and the upper quartile (Fig. 3A) and regarding the median in the case of the intercept (Fig. 3B). Therefore, the RF approaches with the different combination of predictors seemed the most promising. $RF_{sp\text{-}ele}$ was the only combination that gave the most promising results for the slope and intercept as well. This was not true for the other options, because those gave incoherent errors: for some the squared errors were higher for the slope while indicating relatively smaller errors for the intercept, and vice-versa (Fig. 3). Thus, to be consistent, the random forest approach with buffer distance, and with the inclusion of elevation as predictors ($RF_{sp\text{-}ele}$) was chosen to provide the best interpolation for the slope (Fig. 2A) and the intercept (Fig. 2B) of the Mediterranean LWML.

3.3. Spatial patterns and regional differences

Comparing not only the estimated point values (Sect. 3.2) but the spatial trend seen in the spatially continuous interpolated products

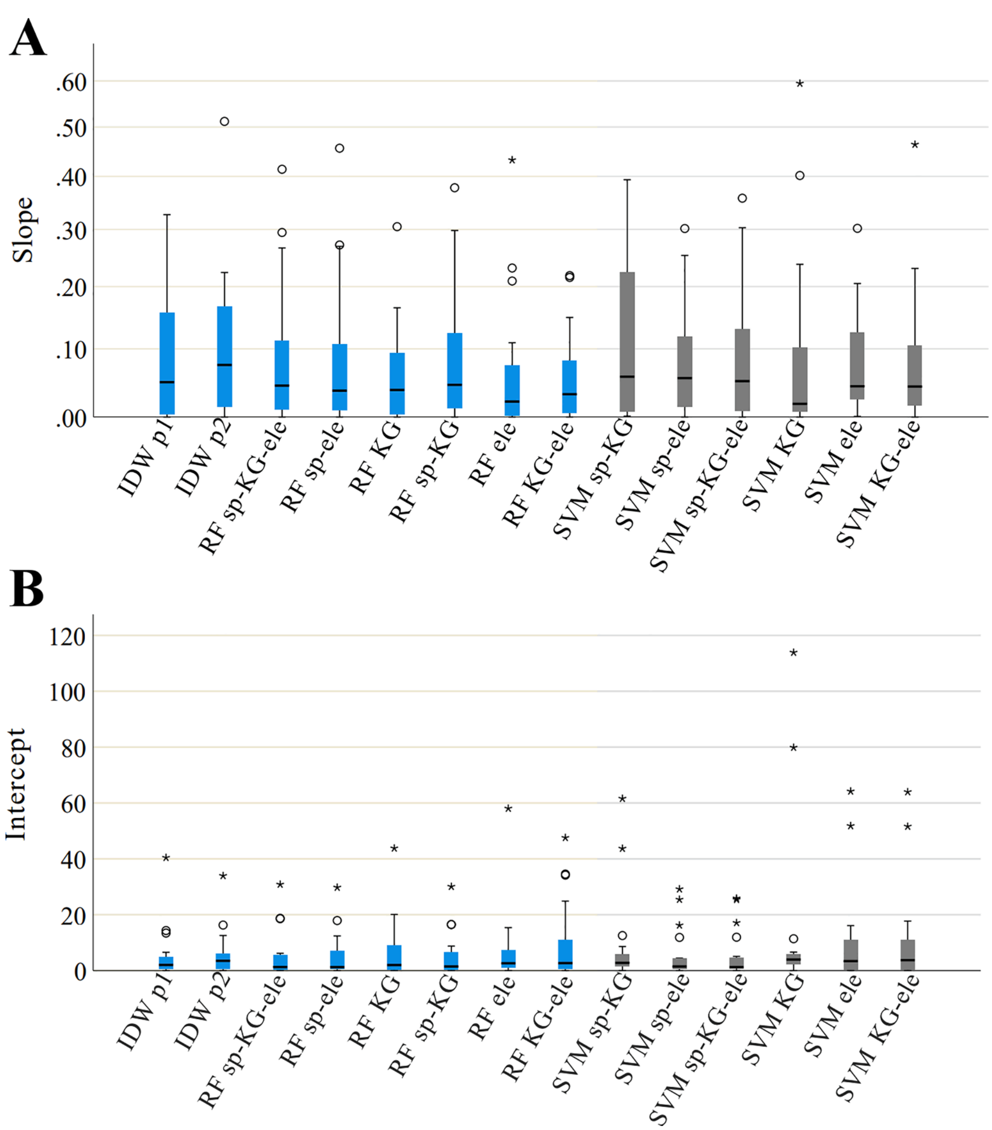


Fig. 3. Box-and-whiskers plots of the squared errors. Boxes represent the squared errors between the measured values of both sets of validation stations (Vs1 and Vs2) and the values for slope A) and the intercept B) estimated at their locations. The boxes indicate the interquartile interval, the black line in the middle the median, while the values outside 1.5-times the interquartile interval are indicated by a circle and the ones with higher values than 3 times the interquartile range are extreme values indicated with an asterisk (Kovács et al., 2012). Grey bars represent the errors of the SVM results omitted in the first round of the analysis, see text for details.

(Fig. 2 and Figs. S1 and S2), it was observed that these concur with the findings documented in numerous studies from the region over the decades. There are, however, some differences that are explained later in this section.

A recent regional study described the LMWL across NW Iberia on the basis of a small regional monitoring network (Moreno et al., 2021) partially overlapping with the chosen timeframe (2010–2012), but not meeting the quality criteria described in Sect. 2.3.1. Apart from Barcelona and Mallorca, the stations used in that study were not included in the present analysis, hence offering the possibility of an independent comparison to assess the interpolated LMWL parameters. Along the Cantabrian Coast, the difference in the slope between the presented model and a reference study (Moreno et al., 2021) is 0.6. This may be because the region is on the border of the present study area, where the slope values were merely extrapolated. In contrast to this, the slope does represent quite a good match with the other stations (Moreno et al., 2021), with an average difference of 0.12. In an Iberian regional reference study, the intercept ranges between 1 and 5‰ (Moreno et al., 2021), which is similar to the values in the present estimations (Fig. 2. B).

Heading east, the regional trend previously documented across the Apennine Peninsula (Giustini et al., 2016) is in harmony with the pattern documented here: a continuous decrease in slope (from 8 to 7) and intercept (from 9.5 to 6‰) from north to south (Fig. 2A). Specifically, the slope parameter displays a good match over the whole region (Fig. 2A), while the values of the intercept are smaller (Fig. 2B) compared to those in a reference study (Giustini et al., 2016). If the present estimations are compared with the LMWL of a more localized study from NE Italy (Masiol et al., 2021), the slope values show a good match, while the intercept estimated in the present work is slightly higher by ~1‰. On the SE coast of the Gulf of Genoa (Fig. 1A), a regional study estimated a slope of 7.28 and intercept: 7.3‰ (Natali et al., 2021), which gave a rather good match with the presented modelled values (Fig. 2). It should be noted, however, that two out of the four central Italian stations were used in the interpolation exercise after preprocessing. Therefore, this comparison cannot be taken as fully independent. Similar slope and intercept values are observed across Sicily compared to the sub-regional outcrop of the model for the Apennine Peninsula (Giustini et al., 2016), and a better match is seen in the highly elevated NW coastal stations of Sicily when compared to a sub-regional model of the island (Liotta et al., 2006). Compared to a study assessing the regional LMWLs in Sicily, the presented slope values are slightly higher and so are the intercept values (~1‰) (Liotta et al., 2013), although this difference might come from the relatively short period covered (Jul 2004 – Jun 2006), which does not meet the necessary “four year criterion” (Putman et al., 2019).

An independent dataset from eastern Serbia incorporating three stations overlaps with the study period (Vasić, 2017). The specific locations of the stations used in that reference work were not documented and the temporal coverage was not continuous, thus these could only be used with reservations for comparison and are not included in the modeling. The slope of individual stations ($n = 3$) ranged from 6.6 to 7.2, with the intercept spanning −2.4 to 4.8‰ (Vasić, 2017). The present model provided a spatially much smoother and somewhat higher estimation for the slope and intercept for the corresponding area, with values of ~7.4 and ~5.3‰, respectively (Fig. 2), probably due to their applied method (OLS) and temporal gaps and in the reference data (Vasić, 2017).

Considering the southern sector of the Balkan Peninsula, reported LMWLs display an unusually large scattering in slope and intercept (from 5.5 to 9.2 and from −1.2 to 21‰, respectively), even for locations close to each other in southern Greece (Dotsika et al., 2010). On the contrary, the macro-regional Mediterranean model presented here displays much smoother, and hence realistic, estimates of the LMWL parameters. Heading further east, the slope and intercept indicate a rather meridional trend across Greece. The slope gradually decreases from ~7.3 to ~6.9 going west to east across Greece (Fig. 2A), while the intercept decreases from the Albanian-Greek border (~5.9‰) to < 3‰ also towards Turkey (Fig. 2B).

For the final subregional comparison, in Cyprus, a LMWL ($\delta^2H = 6.58 \pm 0.13 \times \delta^{18}O + 12.64 \pm 0.91$) was derived in a recent study from the mean precipitation-weighted values of 16 monitoring stations (Christofi et al., 2020), updating older findings based on many fewer stations, $\delta^2H = 5.75 \times \delta^{18}O + 3.6$ (Jacovides, 1979). Taking the previous estimations into account, the present model's slope for the area (~6.8) agrees well with the most recent one, while the intercept (~7‰) is below the recent estimate, but definitely higher than the older study's intercept value.

Regarding North Africa, a study combining rainfall events and monthly data from 17 monitoring stations defined two LMWLs (Ait Brahim et al. (2016): Fig. 5 therein), the estimated parameters of which are both higher than those yielded by the present estimation (slope ~7; intercept ~7.5‰). This may well be due to relative scarcity of station data in this subregion passing the quality control requirements (Sect. 2.3.1) of the present study.

3.4. Regional meteoric water lines in the eastern and western Mediterranean domain?

In the eastern sector of the region the current macro-regional Mediterranean model (i) displays a slightly higher slope (around and between 6.8 and 6.9) compared to previous regional works for precipitation from the 1960s (< 6.3; (Gat and Dansgaard, 1972)), but (ii) better agrees if all rain events from the 1990s are considered (slope < 8; (Ayalon et al., 1998)). The present model's highest intercept values scatter around 15‰ in the Eastern Mediterranean sector. Thus, neither the slope, nor the intercept reproduce the frequently cited benchmark values of the $\delta^2H = 8 \times \delta^{18}O + 22$ EMMWL ‘variant’. It should be noted that both parameters show a quite homogeneous pattern in this subregion, supporting the derivation of a regional MWL. The EMMWL may well have proven itself to be a useful isotope-hydrological benchmark, but cannot be taken as an actual empirical relationship validly describing the meteoric waters in the eastern Mediterranean following the fundamental rules of deriving meteoric water line(s) (Gat, 2005).

In the western Mediterranean the presence of a regional MWL has been suggested, with a slope of 8 and intercept 13.7‰ (Celle-Jeanton et al., 2001) overarching a vast part of the area. However, the present model – based on a much higher number of stations – shows clear differences across the domain under consideration (Fig. 2). The slope varies between ~6.5 and ~7.5, with the lowest values in Algeria, and the highest in Iberia (Fig. 2A), while the intercept ranges from ~2.5 to ~6.5‰, with the lowest values seen again in Algeria, and the highest values near the Gulf of Genoa (Fig. 2B). The spatial heterogeneity of LMWL parameters based on precipitation from between 2000 and 2015 in the western Mediterranean Basin does not support the definition of a uniform regional MWL for this region. Further use of the “Western Mediterranean Meteoric Water Line” as an isotope hydrological benchmark therefore seems to be untenable. Instead, the use of (i) the local estimates from nearby stations with long-term monitoring, or in the absence of a nearby station, (ii) estimates retrieved from the current interpolated product is recommended (Fig. 2).

4. Conclusions

In the study the long-term precipitation stable isotopic data of 249 Mediterranean precipitation monitoring stations were assessed to derive a revised and extended spatially continuous geostatistical model of the local meteoric water lines of the region for the early twenty-first century. In the chosen focus period (2000–2015) rigorous preprocessing consisted of iteratively assessing (i) the local indicator of spatial association together with (ii) d-excess values to find outliers. As a result, 62 stations were retained with at least four years of data coverage ensuring a quality-controlled dataset for the derivation of the LMWLs with RMA regression and the interpolation of the slope and intercept of the δ^2H – δ^{18}O relationship across the region using random forest interpolation with buffer distance. The slope and intercept of the LMWLs ranged from ~5.9 to 8.2 and −3.9 to 16.1‰, respectively.

The macro-regional Mediterranean model of δ^2H – δ^{18}O covariance presented here provides spatially continuous predictions of the slope and intercept of the LMWL of modern (post-2000 CE) precipitation, updating the results of subregional studies from the twentieth century. The model provides an opportunity to revisit the classical concepts of the EMMWL and the WMMWL. With regard to the former, the coherent pattern of rather high intercept values in the eastern Mediterranean reinforces the validity and utility of the concept of the EMMWL for the 21st century, too, but in the west there was no homogeneous pattern of LMWL parameters. On the basis of the results of the present model, it is therefore suggested that the WMMWL should be abandoned as an isotope-hydrological benchmark for modern precipitation. The sub-regional to regional differences observed in the interpolated LMWLs highlight the importance of deriving fine scale spatially continuous estimations of the MWL and not aggregating data together from numerous stations that may be hundreds or even thousands of kilometers apart. It is therefore advised that either local estimates from close stations with long-term monitoring, or in the absence of a nearby station, estimates retrieved from the current interpolated product (Figs. S1 and S2) should be used for the region.

Author contributions

Z.K conceived and designed the study with contribution from I.G.H. I.G.H, D.E, A.E.S., and G.Sz performed the analysis. D.E. and I.G.H produced the figures. I.G.H and Z.K wrote the paper, with contributions from D.E., A.E.S., G.Sz. and P.V. The FLAE approach was applied to the sequence of authors; see https://doi.org/10.1371/journal.pbio.0050018. All authors have read and agreed to the published version of the manuscript.

Funding

This research was supported by the National Research, Development and Innovation Office under Grants SNN118205 and by the Slovenian Research Agency ARRS under Grants N1-0054 and P1-0143.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found at: https://nucleus.iaea.org/wiser/index.aspx and in the papers cited, see Table S1 and Sect. 2.2 for further details.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The source of the used data is explained in the Data Availability Statement in the MS.

Acknowledgments

The authors thank Paul Thatcher for his work on the English version; the authors are grateful for Christof Christofi, Jasper Wassenburg, Enrico Selmo and Roxana Ionete for providing data. The results of this study have been discussed within the COST Action: "WATSON" CA19120.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2022.128925.

References

Aggarwal, P.K., Gat, J.R., Froehlich, K.F.O., 2005. Isotopes in the Water Cycle: Past. Present and Future of a Developing Science, Springer, Netherlands, Dordrecht, p. 382.

Ait Brahim, Y., et al., 2016. Elucidating the climate and topographic controls on stable isotope composition of meteoric waters in Morocco, using station-based and spatially-interpolated data. J. Hydrol. 543, 305–315. https://doi.org/10.1016/j.jhydrol.2016.10.001.

Anselin, L., 1996. The Moran scatterplot as an ESDA tool to assess local instability in spatial association. In: Fischer, M., Scholten, H.J., Unwin, D. (Eds.), Spatial Analytical Perspectives on GIS Routledge, London, UK, pp. 111-126.

Argiriou, A.A., Lykoudis, S., 2006. Isotopic composition of precipitation in Greece. J. Hydrol. 327 (3), 486–495. https://doi.org/10.1016/j.jhydrol.2005.11.053.

Ayalon, A., Bar-Matthews, M., Sass, E., 1998. Rainfall-recharge relationships within a karstic terrain in the Eastern Mediterranean semi-arid region, Israel: δ^{18}O and δD characteristics. J. Hydrol. 207 (1), 18–31. https://doi.org/10.1016/S0022-1694(98)00119-X.

Bajjali, W., Abu-Jaber, N., 2001. Climatological signals of the paleogroundwater in Jordan. J. Hydrol. 243 (1), 133–147. https://doi.org/10.1016/S0022-1694(00)00409-1.

Bartholy, J., Pongrácz, R., Pattantyús-Ábrahám, M., 2009. Analyzing the genesis, intensity, and tracks of western Mediterranean cyclones. Theor. Appl. Climatol. 96 (1), 133–144. https://doi.org/10.1007/s00704-008-0082-9.

Biau, G., Scornet, E., 2016. A random forest guided tour. TEST 25 (2), 197–227. https://doi.org/10.1007/s11749-016-0481-7.

Bojar, A.-V., Halas, S., Bojar, H.-P., Chmiel, S., 2017. Stable isotope hydrology of precipitation and groundwater of a region with high continentality, South Carpathians, Romania. Carpathian J. Earth Environ. Sci. 12 (2), 513–524.

Boschetti, T., Cifuentes, J., Iacumin, P., Selmo, E., 2019. Local meteoric water line of Northern Chile (18° S–30° S): an application of error-in-variables regression to the oxygen and hydrogen stable isotope ratio of precipitation. Water 11 (4), 791.

Bottyán, E., Czuppon, G., Weidinger, T., Haszpra, L., Kármán, K., 2017. Moisture source diagnostics and isotope characteristics for precipitation in east Hungary: implications for their relationship. Hydrol. Sci. J. 62 (12), 2049–2060. https://doi.org/10.1080/02626667.2017.1358450.

Boumaiza, L., Chesnaux, R., Drias, T., Walter, J., Stumpp, C., 2021. Using vadose-zone water stable isotope profiles for assessing groundwater recharge under different climatic conditions. Hydrol. Sci. J. 66 (10), 1597–1609. https://doi.org/10.1080/02626667.2021.1957479.

Bowen, G.J., 2008. Spatial analysis of the intra-annual variation of precipitation isotope ratios and its climatological corollaries. J. Geophys. Res. Atmos. 113 (D5) https://doi.org/10.1029/2007JD009295.

Bowen, G.J., Cai, Z., Fiorella, R.P., Putman, A.L., 2019. Isotopes in the water cycle: regional- to global-scale patterns and applications. Annu. Rev. Earth Planet. Sci. 47 (1), 453–479. https://doi.org/10.1146/annurev-earth-053018-060220.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32. https://doi.org/10.1023/A:1010933404324.

Brkić, Ž., Kuhta, M., Hunjak, T., Larva, O., 2020. Regional isotopic signatures of groundwater in Croatia. Water 12 (7), 1983.

Carroll, R.J., Ruppert, D., 1996. The use and misuse of orthogonal regression in linear errors-in-variables models. Am. Statist. 50 (1), 1–6. https://doi.org/10.1080/00031305.1996.10473533.

Celle-Jeanton, H., Travi, Y., Blavoux, B., 2001. Isotopic typology of the precipitation in the Western Mediterranean Region at three different time scales. Geophys. Res. Lett. 28 (7), 1215–1218. https://doi.org/10.1029/2000GL012407.

Cervi, F., et al., 2017. Isotopic features of precipitation and groundwater from the Eastern Alps of Italy: results from the Mt. Tinisa hydrogeological system. Environ. Earth Sci. 76 (12), 410. https://doi.org/10.1007/s12665-017-6748-9.

Christofi, C., Bruggeman, A., Kuells, C., Constantinou, C., 2020. Isotope hydrology and hydrogeochemical modeling of Troodos Fractured Aquifer, Cyprus: The development of hydrogeological descriptions of observed water types. Appl. Geochem. 123, 104780 https://doi.org/10.1016/j.apgeochem.2020.104780.

Ciric, D., Nieto, R., Losada, L., Drumond, A., Gimeno, L., 2018. The mediterranean moisture contribution to climatological and extreme monthly continental precipitation. Water 10 (4).

Clark, I.D., Fritz, P., 1997. Environmental Isotopes in Hydrogeology. Taylor & Francis.

Clarke, M.R.B., 1980. The reduced major axis of a bivariate sample. Biometrika 67 (2), 441–446. https://doi.org/10.1093/biomet/67.2.441.

Cook, B.I., Anchukaitis, K.J., Touchan, R., Meko, D.M., Cook, E.R., 2016. Spatiotemporal drought variability in the Mediterranean over the last 900 years. J. Geophys. Res. Atmos. 121 (5), 2060–2074. https://doi.org/10.1002/2015JD023929.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20 (3), 273–297. https://doi.org/10.1007/BF00994018.

Craig, H., 1961. Isotopic variations in meteoric waters. Science 133 (3465), 1702–1703. https://doi.org/10.1126/science.133.3465.1702.

Crawford, J., Hughes, C.E., Lykoudis, S., 2014. Alternative least squares methods for determining the meteoric water line, demonstrated using GNIP data. J. Hydrol. 519, 2331–2340. https://doi.org/10.1016/j.jhydrol.2014.10.033.

Dansgaard, W., 1964. Stable isotopes in precipitation. Tellus 16, 436–468.
Dotsika, E., Lykoudis, S., Poutoukis, D., 2010. Spatial distribution of the isotopic composition of precipitation and spring water in Greece. Global Planet. Change 71 (3), 141–149. https://doi.org/10.1016/j.gloplacha.2009.10.007.
El Ouali, A., et al., 2022. Isotopic characterization of rainwater for the development of a local meteoric water line in an arid climate: the Case of the Wadi Ziz Watershed (South-Eastern Morocco). Water 14 (5), 779.
El-Asrag, A., 2005. Effect of synoptic and climatic situations on fractionation of stable isotopes in rainwater over Egypt and east Mediterranean. Isotopic composition of precipitation in the Mediterranean Basin in relation to air circulation patterns and climate 51–73.
Elghawi, R., Pekhazis, K., Doummar, J., 2021. Multi-regression analysis between stable isotope composition and hydrochemical parameters in karst springs to provide insights into groundwater origin and subsurface processes: regional application to Lebanon. Environ. Earth Sci. 80 (6), 230. https://doi.org/10.1007/s12665-021-09519-4.
Erdélyi, D., et al., 2023. Predicting spatial distribution of stable isotopes in precipitation by classical geostatistical- and machine learning methods. J. Hydrol., 129129 https://doi.org/10.1016/j.jhydrol.2023.129129.
Fernández-Chacón, F., Benavente, J., Rubio-Campos, J.C., Kohfahl, C., Jiménez, J., Meyer, H., Hubberten, H., Pekdeger, A., 2010. Isotopic composition (δ^{18}O and δD) of precipitation and groundwater in a semi-arid, mountainous area (Guadiana Menor basin, Southeast Spain). Hydrol. Process. 24 (10), 1343–1356. https://doi.org/10.1002/hyp.7597.
Fischer, M.J., Mattey, D., 2012. Climate variability and precipitation isotope relationships in the Mediterranean region. J. Geophys. Res. Atmos. 117 (D20) https://doi.org/10.1029/2012JD018010.
Fórizs, I., 2003. Isotopes as natural tracers in the watercycle: examples from the Carpathian Basin. Studia UBB Physica 1 (48), 69–77.
Fórizs, I., Kern, Z., Csicsák, J., Csurgó, G., 2020. Monthly data of stable isotopic composition (δ^{18}O, δ^2H) and tritium activity in precipitation from 2004 to 2017 in the Mecsek Hills, Hungary. Data in Brief 32, 106206. https://doi.org/10.1016/j.dib.2020.106206.
García-Ruiz, J.M., López-Moreno, J.I., Vicente-Serrano, S.M., Lasanta-Martínez, T., Beguería, S., 2011. Mediterranean water resources in a global change scenario. Earth Sci. Rev. 105 (3), 121–139. https://doi.org/10.1016/j.earscirev.2011.01.006.
Gat, J.R., 1996. Oxygen and hydrogen isotopes in the hydrologic cycle. Annu. Rev. Earth Planet. Sci. 24 (1), 225–262. https://doi.org/10.1146/annurev.earth.24.1.225.
Gat, J.R., 2005. Some Classical Concepts of Isotope Hydrology. In: Aggarwal, P.K., Gat, J. R., Froehlich, K.F.O. (Eds.), Isotopes in the Water Cycle: Past, Present and Future of a Developing Science. Springer Netherlands, Dordrecht, pp. 127-137. DOI:10.1007/1-4020-3023-1_10.
Gat, J.R., Carmi, I., 1970. Evolution of the isotopic composition of atmospheric waters in the Mediterranean Sea area. J. Geophys. Res. (1896–1977), 75(15): 3039-3048. Doi: https://doi.org/10.1029/JC075i015p03039.
Gat, J.R., Dansgaard, W., 1972. Stable isotope survey of the fresh water occurrences in Israel and the Northern Jordan Rift Valley. J. Hydrol. 16 (3), 177–211. https://doi.org/10.1016/0022-1694(72)90052-2.
Gat, J.R., Mook, W.G., Meijer, H.A., 2001. Environmental isotopes in the hydrological cycle, 2. International Atomic Energy Agency, Paris, 73 pp.
Giustini, F., Brilli, M., Patera, A., 2016. Mapping oxygen stable isotopes of precipitation in Italy. J. Hydrol.: Reg. Stud. 8, 162–181. https://doi.org/10.1016/j.ejrh.2016.04.001.
Golobočanin, D., Ogrinc, N., Bondzić, A., Miljević, N., 2007. Isotopic characteristics of meteoric waters in the Belgrade region. Isot. Environ. Health Stud. 43 (4), 355–367. https://doi.org/10.1080/10256010701702929.
Hatvani, I.G., Erdélyi, D., Vreča, P., Kern, Z., 2020. Analysis of the spatial distribution of stable oxygen and hydrogen isotopes in precipitation across the Iberian Peninsula. Water 12 (2), 481. https://doi.org/10.3390/w12020481.
Hengl, T., Heuvelink, G.B.M., Stein, A., 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma 120 (1), 75–93. https://doi.org/10.1016/j.geoderma.2003.08.018.
Hoerling, M., et al., 2012. On the increased frequency of Mediterranean drought. J. Clim. 25 (6), 2146–2161. https://doi.org/10.1175/jcli-d-11-00296.1.
Hssaisoune, M., et al., 2022. Isotopic and chemical tracing for residence time and recharge mechanisms of groundwater under semi-arid climate: case from Rif mountains (Northern Morocco). Geosciences 12 (2), 74.
Hunjak, T., Lutz, H.O., Roller-Lutz, Z., 2013. Stable isotope composition of the meteoric precipitation in Croatia. Isot. Environ. Health Stud. 49 (3), 336–345. https://doi.org/10.1080/10256016.2013.816697.
IAEA, 1992. Statistical treatment of data on environmental isotopes in precipitation. Technical Report Series International Atomic Energy Agency, Vienna, 781place pp.
IAEA, 2005. Isotopic composition of precipitation in the Mediterranean Basin in relation to air circulation patterns and climate. TECDOC Series. International Atomic Energy Agency, Vienna, Austria, 223 pp.
Iglesias, A., Garrote, L., Flores, F., Moneo, M., 2007. Challenges to manage the risk of water scarcity and climate change in the Mediterranean. Water Resour. Manag. 21 (5), 775–788. https://doi.org/10.1007/s11269-006-9111-6.
Ishwaran, H., Kogalur, U., Kogalur, M., 2021. RandomForestSRC: Fast Unified Random Forests for Survival. Regression, and Classification (RF-SRC) [(accessed on 16 July 2020)].
Ishwaran, H., Kogalur, U.B., Blackstone, E.H., Lauer, M.S., 2008. Random survival forests. Annals Appl. Stat., 2(3): 841-860, 20.
Jacovides, J., 1979. Environmental isotope survey (Cyprus). Final report on IAEA, research contract No: 1039, Ministry of Agriculture and Natural Resources, Department of Water Development, Nicosia, Cyprus.
Kattan, Z., 1997. Chemical and environmental isotope study of precipitation in Syria. J. Arid Environ. 35 (4), 601–615. https://doi.org/10.1006/jare.1996.0228.
Kelemen, F.D., Bartholy, J., Pongracz, R., 2015. Multivariable cyclone analysis in the Mediterranean region. Időjárás 119 (2), 159–184.
Kelley, C.P., Mohtadi, S., Cane, M.A., Seager, R., Kushnir, Y., 2015. Climate change in the Fertile Crescent and implications of the recent Syrian drought. Proc. Natl. Acad. Sci. 112 (11), 3241–3246. https://doi.org/10.1073/pnas.1421533112.
Kendall, C., Coplen, T.B., 2001. Distribution of oxygen-18 and deuterium in river waters across the United States. Hydrol. Process. 15 (7), 1363–1393. https://doi.org/10.1002/hyp.217.
Koeniger, P., Margane, A., 2014. Stable isotope investigations in the Jeita Spring catchment.
Koeniger, P., Margane, A., Abi-Rizk, J., Himmelsbach, T., 2017. Stable isotope-based mean catchment altitudes of springs in the Lebanon Mountains. Hydrol. Process. 31 (21), 3708–3718. https://doi.org/10.1002/hyp.11291.
Kottek, M.G., Jürgen; Beck, Christoph; Rudolf, Bruno; Rubel, Franz, 2006. World Map of the Köppen-Geiger climate classification updated. Meteorologische Zeitschrift, 15 (3): 259-263. DOI:10.1127/0941-2948/2006/0130.
Kovács, J. et al., 2012. Analysis of Water Quality Data for Scientists. In: Kostas Voudouris, Voutsa, D. (Eds.), Water Quality Monitoring and Assessment. InTech, pp. 65-94. DOI:10.5772/32173.
Krajcar Bronić, I., et al., 2020a. Long-term isotope records of precipitation in Zagreb, Croatia. Water 12 (1), 226.
Krajcar Bronić, I., et al., 2020b. Isotope Composition of precipitation, groundwater, and surface and lake waters from the Plitvice Lakes, Croatia. Water 12 (9), 2414.
Kruskal, W.H., Wallis, W.A., 1952. Use of ranks in one-criterion variance analysis. J. Am. Stat. Assoc. 47 (260), 583–621. https://doi.org/10.2307/2280779.
Kuhn, M. et al., 2020. caret: Classification and Regression Training. R package version 6.0-86. Astrophysics Source Code Library: Cambridge, MA, USA.
Lécuyer, C., Bojar, A.-V., Daux, V., Legendre, S., 2020. Geographic variations in the slope of the δ^2H–δ^{18}O meteoric water line over Europe: a record of increasing continentality. Geological Society, London, Special Publications, 507: SP507-2020-68. DOI:10.1144/sp507-2020-68.
Legendre, P., 2018. lmodel2: Model II Regression. R package version 1.7-3.
Li, J., Heap, A.D., Potter, A., Daniell, J.J., 2011. Application of machine learning methods to spatial interpolation of environmental variables. Environ. Model. Softw. 26 (12), 1647–1659. https://doi.org/10.1016/j.envsoft.2011.07.004.
Liotta, M., et al., 2013. Isotopic composition of precipitation and groundwater in Sicily, Italy. Appl. Geochem. 34, 199–206. https://doi.org/10.1016/j.apgeochem.2013.03.012.
Liotta, M., Favara, R., Valenza, M., 2006. Isotopic composition of the precipitations in the central Mediterranean: origin marks and orographic precipitation effects. J. Geophys. Res. Atmos. 111 (D19) https://doi.org/10.1029/2005JD006818.
Liu, Y., Wang, Y., Zhang, J., 2012. New machine learning algorithm: random forest. In: Liu, B., Ma, M., Chang, J. (Eds.), Information Computing and Applications. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 246–252.
Longinelli, A., et al., 2008. A stable isotope study of the Garda Lake, northern Italy: Its hydrological balance. J. Hydrol. 360 (1), 103–116. https://doi.org/10.1016/j.jhydrol.2008.07.020.
Longinelli, A., Anglesio, E., Flora, O., Iacumin, P., Selmo, E., 2006. Isotopic composition of precipitation in Northern Italy: reverse effect of anomalous climatic events. J. Hydrol. 329 (3), 471–476. https://doi.org/10.1016/j.jhydrol.2006.03.002.
Longinelli, A., Selmo, E., 2003. Isotopic composition of precipitation in Italy: a first overall map. J. Hydrol. 270 (1), 75–88. https://doi.org/10.1016/S0022-1694(02)00281-0.
Marchina, C., et al., 2020. Alternative methods to determine the δ^2H-δ^{18}O relationship: An application to different water types. J. Hydrol. 587, 124951 https://doi.org/10.1016/j.jhydrol.2020.124951.
Marković, T., Karlović, I., Perčec Tadić, M., Larva, O., 2020. Application of stable water isotopes to improve conceptual model of alluvial aquifer in the Varaždin Area. Water 12 (2), 379.
Masiol, M., et al., 2021. Spatial distribution and interannual trends of δ^{18}O, δ^2H, and deuterium excess in precipitation across North-Eastern Italy. J. Hydrol. 598, 125749 https://doi.org/10.1016/j.jhydrol.2020.125749.
Meyer, D., Dimitriadou, E., Kurt, H., Andreas, W., Friedrich, L., 2019. e1071: misc functions of the department of statistics, probability theory group (Formerly: E1071), TU Wien. R package version 1. 7–1.
Mood, A.M., 1950. Introduction to the Theory of Statistics.
Moran, P.A.P., 1948. The interpretation of statistical maps. J. Roy. Stat. Soc.: Ser. B (Methodol.) 10 (2), 243–251.
Moreno, A., et al., 2021. Measurement report: Spatial variability of northern Iberian rainfall stable isotope values – investigating atmospheric controls on daily and monthly timescales. Atmos. Chem. Phys. 21 (13), 10159–10177. https://doi.org/10.5194/acp-21-10159-2021.
Natali, S., et al., 2021. Meteorological and geographical control on stable isotopic signature of precipitation in a western Mediterranean area (Tuscany, Italy): Disentangling a complex signal. J. Hydrol. 603, 126944 https://doi.org/10.1016/j.jhydrol.2021.126944.
Nelson, D.B., Basler, D., Kahmen, A., 2021. Precipitation isotope time series predictions from machine learning applied in Europe. Proc. Natl. Acad. Sci., 118(26): e2024107118. DOI:doi:10.1073/pnas.2024107118.
Paar, D., Mance, D., Stroj, A., Pavić, M., 2019. Northern Velebit (Croatia) karst hydrological system: results of a preliminary ^{2}H and ^{18}O stable isotope study. Geol. Croat. 72 (3), 205–213.

Paternoster, M., Liotta, M., Favara, R., 2008. Stable isotope ratios in meteoric recharge and groundwater at Mt. Vulture volcano, southern Italy. J. Hydrol. 348 (1), 87–97. https://doi.org/10.1016/j.jhydrol.2007.09.038.

Prasad, A.M., Iverson, L.R., Liaw, A., 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems 9 (2), 181–199. https://doi.org/10.1007/s10021-005-0054-1.

Putman, A.L., Fiorella, R.P., Bowen, G.J., Cai, Z., 2019. A global perspective on local meteoric water lines: meta-analytic insight into fundamental controls and practical constraints. Water Resour. Res. 55 (8), 6896–6910. https://doi.org/10.1029/2019WR025181.

R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rindsberger, M., Jaffe, S., Rahamim, S., Gat, J.R., 1990. Patterns of the isotopic composition of precipitation in time and space: data from the Israeli storm water collection program. Tellus B 42 (3), 263–271. https://doi.org/10.1034/j.1600-0889.1990.t01-2-00005.x.

Rozanski, K., Araguás-Araguás, L., Gonfiantini, R., 1993. Isotopic patterns in modern global precipitation. In: Swart, P.K., Lohmann, K.C., Mckenzie, J., Savin, S. (Eds.), Climate Change in Continental Isotopic Records. American Geophysical Union, USA, pp. 1–36. https://doi.org/10.1029/GM078p0001.

Serianz, L., Cerar, S., Vreča, P., 2021. Using stable isotopes and major ions to identify recharge characteristics of the Alpine groundwater-flow dominated Triglavska Bistrica River. Geologija 64 (2), 205–220. https://doi.org/10.5474/geologija.2021.012.

Sharp, Z., 2017. Principles of stable isotope geochemistry.

Surić, M., Lončarić, R., Bočić, N., Lončar, N., Buzjak, N., 2018. Monitoring of selected caves as a prerequisite for the speleothem-based reconstruction of the Quaternary environment in Croatia. Quat. Int. 494, 263–274. https://doi.org/10.1016/j.quaint.2017.06.042.

Tappa, D.J., Kohn, M.J., McNamara, J.P., Benner, S.G., Flores, A.N., 2016. Isotopic composition of precipitation in a topographically steep, seasonally snow-dominated watershed and implications of variations from the global meteoric water line. Hydrol. Process. 30 (24), 4582–4592. https://doi.org/10.1002/hyp.10940.

Trabelsi, R., Matsumoto, T., Zouari, K., Trabelsi, M., Kumar, B., 2020. Investigation of paleoclimate signatures in Sfax deep groundwater (Southeastern Tunisia) using environmental isotopes and noble gases. Quat. Int. 547, 208–219. https://doi.org/10.1016/j.quaint.2019.04.001.

Túri, M., et al., 2019. Tracing groundwater recharge conditions based on environmental isotopes and noble gases, Lom depression, Bulgaria. J. Hydrol: Regional Stud. 24, 100611 https://doi.org/10.1016/j.ejrh.2019.100611.

Túri, M., et al., 2020. Paleotemperature reconstruction using environmental isotopes and noble gases in groundwater in Morocco. Hydrgeol. J. 28 (3), 973–986. https://doi.org/10.1007/s10040-020-02121-1.

Varlam, C., Duliu, O.G., Ionete, R.E., Costinel, D., 2021. Time series analysis of the δ^2H, δ^{18}O values in correlation with monthly temperature, relative humidity and precipitation in Râmnicu Vâlcea, Romania: 2012–2018. Geol. Soc. Lond. Spec. Publ. 507 (1), 77. https://doi.org/10.1144/SP507-2020-56.

Vasić, L.M., 2017. Geneza i uslovi cirkulacije voda kompleksnih karstnih sistema Kučajsko-beljaničkog masiva. University of Belgrade, Belgrade, Serbia, p. 409.

Vasić, L., Palcsu, L., Fen, H., 2019. Groundwater gravitational circulation of Karst Veliko Vrelo and Malo Vrelo springs by isotope and the noble gas method: case study of the Beljanica Massif. Environ. Earth Sci. 78 (10), 1–7.

Villani, L., Castelli, G., Piemontese, L., Penna, D., Bresci, E., 2022. Drought risk assessment in Mediterranean agricultural watersheds: A case study in Central Italy. Agric Water Manag 271, 107748. https://doi.org/10.1016/j.agwat.2022.107748.

Vreča, P., Bronić, I.K., Horvatinčić, N., Barešić, J., 2006. Isotopic characteristics of precipitation in Slovenia and Croatia: comparison of continental and maritime stations. J. Hydrol. 330 (3), 457–469. https://doi.org/10.1016/j.jhydrol.2006.04.005.

Vreča, P., Brenčič, M., Leis, A., 2007. Comparison of monthly and daily isotopic composition of precipitation in the coastal area of Slovenia. Isot. Environ. Health Stud. 43 (4), 307–321. https://doi.org/10.1080/10256010701702739.

Vreča, P., Krajcar Bronić, I., Leis, A., Brenčić, M., 2008. Isotopic composition of precipitation in Ljubljana (Slovenia). Geologija 51 (2), 169. https://doi.org/10.5474/geologija.2008.018.

Vreča, P., Krajcar Bronić, I., Leis, A., 2011. Isotopic composition of precipitation in Portorož (Slovenia). Geologija 54 (1), 129–137. https://doi.org/10.5474/geologija.2011.010.

Vreča, P., Krajcar Bronić, I., Leis, A., Demšar, M., 2014. Isotopic composition of precipitation at the Station Ljubljana (Reaktor), Slovenia – period 2007–2010. Geologija 57 (2), 217–230. https://doi.org/10.5474/geologija.2014.019.

Vreča, P., Bronić, I.K., Leis, A., 2015. Isotopic composition of precipitation at the station Portorož, Slovenia–period 2007–2010. Geologija 58 (2), 233–246. https://doi.org/10.5474/geologija.2015.019.

Vreča, P., Malenšek, N., 2016. Slovenian Network of Isotopes in Precipitation (SLONIP) – a review of activities in the period 1981–2015. Geologija 59 (1), 67–84. https://doi.org/10.5474/geologija.2016.004.

Vreča, P., Pavšek, A., Kocman, D., 2022. SLONIP-A slovenian web-based interactive research platform on water isotopes in precipitation. Water 14 (13), 2127. https://doi.org/10.3390/w14132127.

Wassenburg, J.A., et al., 2016. Reorganization of the North Atlantic Oscillation during early Holocene deglaciation. Nat. Geosci., advance online publication. https://doi.org/10.1038/ngeo2767.

Webster, R., Oliver, M.A., 2008. Geostatistics for Environmental Scientists. Geostatistics for Environmental Scientists. John Wiley & Sons Ltd 330 pp. https://doi.org/10.1002/9780470517277.index.

Wilcox, R.R., 2003. 3 - SUMMARIZING DATA. In: Wilcox, R.R. (Ed.), Applying Contemporary Statistical Techniques. Academic Press, Burlington, pp. 55-91. DOI: https://doi.org/10.1016/B978-012751541-0/50024-9.

Wright, M.N., Ziegler, A., 2015. ranger: A fast implementation of random forests for high dimensional data in C++ and R. arXiv preprint arXiv:1508.04409.

Yoshimura, K., 2015. Stable water isotopes in climatology, meteorology, and hydrology: A review. J. Meteorol. Soc. Japan. Ser. II 93 (5), 513–533. https://doi.org/10.2151/jmsj.2015-036.

Yüce, G., 2005. Determination of the recharge area and salinization degree of karst springs in the Lamas Basin (Turkey). Isot. Environ. Health Stud. 41 (4), 391–404. https://doi.org/10.1080/10256010500384747.

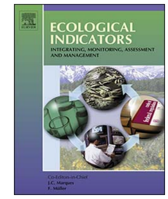
istvan60_104_23

Ecological Indicators 83 (2017) 21–31

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind



Original Articles

Periodic signals of climatic variables and water quality in a river – eutrophic pond – wetland cascade ecosystem tracked by wavelet coherence analysis

István Gábor Hatvani[a,*], Adrienne Clement[b], János Korponai[c,d], Zoltán Kern[a], József Kovács[e]

[a] Institute for Geological and Geochemical Research, Research Centre for Astronomy and Earth Sciences, Hungarian Academy of Sciences, H-1112 Budapest, Budaörsi út 45, Hungary
[b] Budapest University of Technology and Economics, Department of Sanitary and Environmental Engineering, H-1111 Budapest, Műegyetem rakpart 3, Hungary
[c] Eötvös Loránd University, Savaria Campus, Department of Biology, H-9700 Szombathely, Károly Gáspár Tér 4., Hungary
[d] MTA – PE Limnoecology Research Group, University of Pannonia, Warta Vince u. 1, 8200 Veszprém, Hungary
[e] Eötvös Loránd University, Department of Physical and Applied Geology, H-1117 Budapest, Pázmány Péter stny. 1/C, Hungary

ARTICLE INFO

Keywords:
Ecosystem management
Eutrophication
Kis-Balaton water protection system
Macrophyte cover
Meteorological driving effect
Nutrient retention
Vulnerability

ABSTRACT

Lakes are sensitive to changes in their environmental boundary conditions that can be indicated in the periodic behavior of water quality variables. The present work aims to assess the degree to which common annual periodic behavior is present (1994–2010) in the meteorological parameters (global radiation, air temperature, cloud cover), streamflow; and five primary nutrients (e.g. total phosphorus, nitrate-nitrogen) as possible indicators of ecosystem vulnerability in four different ecosystems using wavelet coherence analysis. The cascade system is located in the mouth of a shallow river where the water flows through a eutrophic pond then a disturbed/undisturbed macrophyte covered wetland reaching a large shallow lake. The results highlight the differing abilities of the elements of the cascade of ecosystems to follow seasonality. The changes in water quality (nutrient cycle) in the eutrophic pond most closely mirror meteorological seasonality. The vulnerability of the wetland ecosystem was expressed by its decreased capacity to follow seasonal changes due to high algae loads and additional inflows. Moreover, the wetland proved to be weak and unstable regarding phosphorus and nitrogen retention. With the successful application of wavelet coherence analysis to the "black-box" cascade system the study sets an example for the implications of the method in such combined or stand-alone natural/partially-constructed ecosystems.

1. Introduction

Water, and especially fresh water, is one of the most critical natural resources which is highly endangered by climate change and anthropogenic activity (Vörösmarty et al., 2000). It has been documented that environmental (Reynolds, 1984) and anthropogenic factors (Kovács et al., 2010) govern and may indeed corrupt the capacity of freshwater ecosystems to follow seasonal changes. In the moderate climate zone aquatic ecosystems, e.g. rivers (Wong et al., 1978) and shallow lakes are per se susceptible to eutrophication (Padisak, 1992), while even constructed wetlands (Kadlec, 1999) tend to follow seasonal changes in hydrometeorology as far as the variables describing their quality and/or quantity are concerned. This phenomenon is mirrored in the seasonal behavior of e.g. runoff (Dettinger and Diaz, 2000), concentrations of nitrogen (Exner-Kittridge et al., 2016) and phosphorus forms (Istvánovics, 1988), or phytoplankton biomass (Reynolds, 1984) through the changing temporal-, light- and hydrologic conditions. In all of these cases, these various characteristics hold vital information about the ecological state of the systems, i.e. of the shallow lakes, rivers, constructed/natural wetlands.

Hitherto, the periodic behavior of a certain water quality variable has usually been studied. There are only a few cases in which sets or groups of parameters, e.g. nutrients, ions, etc. (Kovács et al., 2017), or multiple parameters individually (e.g. chlorophyll-*a*, sodium-, potassium ions, nitrate-nitrogen) (Kovács et al., 2010) have been assessed together to describe the overall capacity of a habitat or several habitats, to follow the seasonal changes. Although the studies cited present a significant and validated picture of the periodic behavior of freshwater ecosystems, they do not directly explore the relationship – that is, the coherence – of the periodic behavior of water quality variables with meteorology. This present study aims to remedy this shortcoming and explore the direct relationship of water quality parameters (mostly inorganic nutrients) with local climate and streamflow in a cascade system consisting of a shallow river, a eutrophic pond and a wetland with both an undisturbed- and disturbed habitats.

* Corresponding author.
E-mail address: hatvaniig@gmail.com (I.G. Hatvani).

http://dx.doi.org/10.1016/j.ecolind.2017.07.018
Received 26 April 2017; Received in revised form 4 July 2017; Accepted 6 July 2017
Available online 27 July 2017

1.1. Study area description

The Kis-Balaton Water Protection System (KBWPS) assessed here functions as a treatment reservoir-wetland system, and was constructed to reduce diffuse nutrient loads reaching Lake Balaton, the largest (surface area approx. 594 km^2) lake in Central Europe. Improving water quality and preserving its good ecological status of the lake is one of the primary goals of European water management (EC, 2000; ICPDR, 2015). The largest tributary to the lake, the River Zala, supplies almost 50% of its water and 35–40% of its nutrient input (Hatvani et al., 2014), therefore significantly affecting its water quality. In the nineteenth century the water level of Lake Balaton and the River Zala was regulated (Lotz, 1988). As a result of this artificial modification, the former wetland areas of Kis-Balaton – located in the Lower Zala Valley – partially dried up and to a great extent became incapable of performing their natural filtering function. Combined with increased agricultural activity (e.g. fertilizer usage) and urbanization (e.g. waste water production) in the course of the 20th century, these changes resulted in the continuous deterioration of Balaton's water quality (Hatvani et al., 2014, 2015; Somlyódy et al., 1983) and occasionally led to considerable economic losses in the tourism sector (Istvánovics et al., 2007). To halt and reverse these negative trends, comprehensive measures for nutrient reduction were taken (Hatvani et al., 2015; Somlyódy et al., 1983) resulting in a 50–60% decrease in biologically available nutrients (Padisák et al., 2006a).

An important part of these measures, the Kis-Balaton Water Protection System (KBWPS) was created in two constructional phases. The remains of the former Kis-Balaton Wetland at the mouth of the River Zala (Fig. 1) were revitalized, and in Phase I, an 18 km^2 reservoir was inundated, commencing operation in 1985. With average depth of ~1 m and a water residence time of approx. 30 days (Hatvani, 2014), this has become an algae-dominated "eutrophic pond" (Fig. 1). In it, summer phytoplankton biomass (chlorophyll-a concentration) exceeds 200 mg m^{-3} and is dominated by cyanobacteria. About 80% of the phosphorus (P) loads are bound in algae and sediment (Mátyás et al., 2003). In 1992 Phase II was put into operation, though up to 2014 only a part of it (16 km^2) was inundated. This area (the "wetland") is covered by macrophytes (Fig. 1). The water residence time here is approximately twice as long as in Phase I (Hatvani, 2014). This "classic wetland" part of the system is covered by reed-dominated macrophytes; euphytoplankton species are therefore scarce, while meroplanktonic species can be found in high number in open water patches (WTWD, 2012).

Since the water coming from the River Zala passes through the different ecosystems (habitats) of the KBWPS and changes into lake water, it is suspected that hydrochemical seasonality (Kolander and Tylkowski, 2008; Tanos et al., 2015) – governed mainly by temperature driving the dynamics of biological processes – will be present/corrupted to a different degree in the various habitats mirroring their local characteristics. This is the particular process that is investigated in the present study with state-of-the-art statistical tools using the key link between hydrochemical seasonality and the periodicity of the water quality parameters.

1.2. Study aims

The specific questions of the study were, how are the differences in behavior (e.g. in nutrient retention) of the connected freshwater ecosystems (shallow river, eutrophic pond and an undisturbed/disturbed wetlands) indicated in the change in common periodicity between the daily measured water quality and the meteorological parameters or streamflow? It is to be expected that by exploring the previously mentioned characteristics a far-reaching overall picture may be obtained of the functioning of the cascade system prevailed by a consistent in/anti-phase coherence. This may serve as an example for the assessment of wetlands ecosystems set up with similar mitigation purposes (Cao et al., 2016; Dunne et al., 2015; Martín et al., 2013; Ni et al., 2016) and be a solid foundation laid down for the wider applicability of the methodology in limnology.

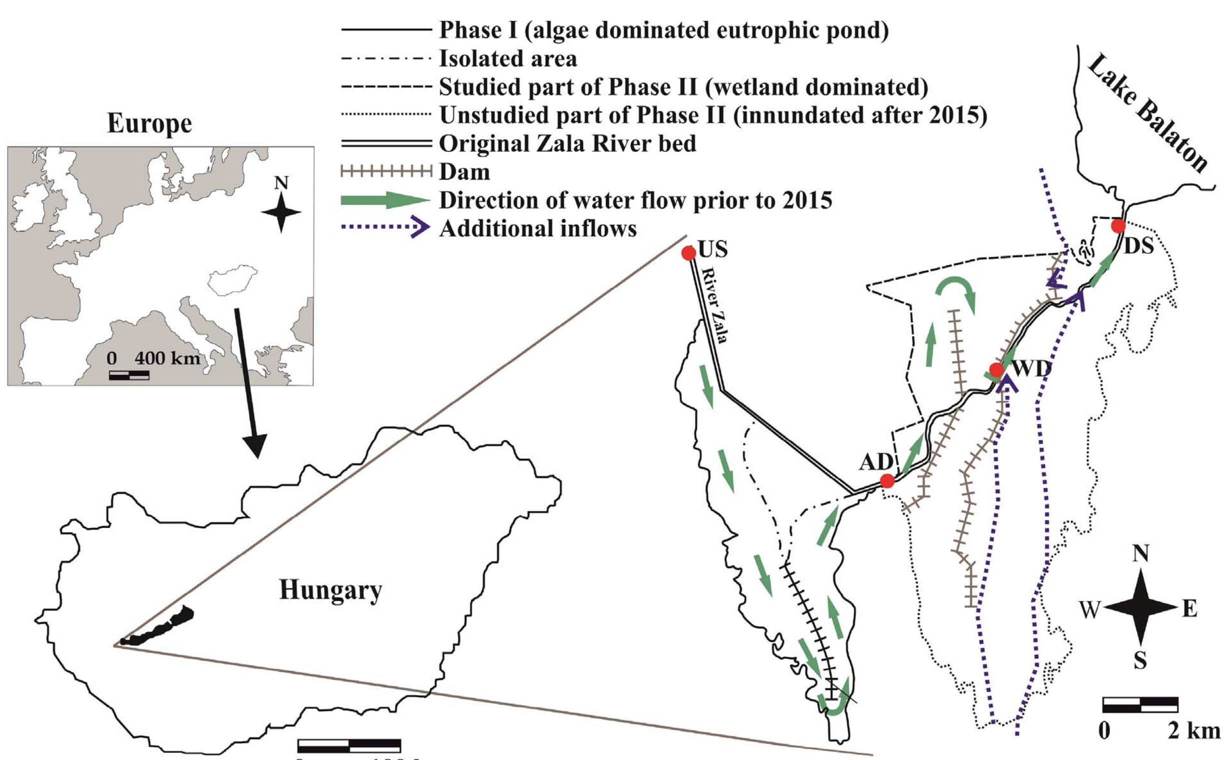


Fig. 1. Location of the study area and the sampling sites (US: upstream, AD: algae dominated, WD: macrophyte dominated, DS: downstream; detailed description in Section 2.2.) marked with red dots (based on Hatvani (2014)). Note, in other studies the sites assessed here (US, AD, WD, DS) are referred to as: Z15, Z11, Kb210, Z27 respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2. Materials and methods

2.1. Dataset used

In the study, the daily time series of 5 water quality parameters (WQPs) – nitrate-nitrogen (NO_3-N); total nitrogen (TN); total phosphorus (TP); phosphate-phosphorus (abbreviated as SRP); total suspended solids (TSS, mg l^{-1}) – were examined, along with background meteorological parameters and daily streamflow (Q; m^3 min^{-1}). This latter is the amount of water passing through a cross-section of the assessed system in a given time. The meteorological parameters included were global radiation (GR, J cm^{-2}), air temperature (T, °C), precipitation (mm) and cloud cover (CC, tenths) (Spinoni et al., 2015). The meteorological parameters together with Q will be referred to as independent variables (IVs) in the study. All data were assessed using wavelet spectrum and wavelet coherence analyses (Torrence and Compo, 1998) for the time interval 1994–2010 from four sampling sites of the KBWPS (Fig. 1). The sites were:

- Upstream, the input of the KBWPS, representing the River Zala, abbreviated in the present study as "US".
- The outflow of the algae-dominated shallow eutrophic pond Phase I, abbreviated in the present study as "AD".
- The outflow of the macrophyte-dominated wetland habitat, representing the undisturbed wetland, abbreviated in the present study as "WD".
- The downstream outlet of KBWPS, including the outflow water of the wetland and additional external inputs reach the system bringing a 40% excess in streamflow (Hatvani et al., 2014), thus representing a "mixed" wetland habitat (disturbed wetland); abbreviated in the present study as "DS".

The latter two (WD and DS) will be referred to together in certain places of the paper as Phase II (Fig. 1). Please note that for Q at WD, the data was only available from 01.01.1995.

2.2. Methodology

The periodic behavior of the independent variables was evaluated using wavelet spectrum analysis to identify those time intervals lacking annual periodicity. Than to find the direct common periodic signal between the water quality parameters and the independent variables, wavelet transform coherence (WTC) was used, as it was applied e.g. to uncover the relationship between climate indices and streamflow variability (Nalley et al., 2016), to explore the relationship between water levels and chlorophyll-a in Lake Baiyangdian (Wang et al., 2012). This approach was also used, e.g. on stable isotopes in precipitation and temperature (Salamalikis et al., 2016), on speleothems and climate variables (Hatvani et al., 2017), or in assessing low-frequency variability in hydroclimate records from east Central Europe (Sen and Kern, 2016).

Wavelet spectrum analysis is considered as a function localized in both frequency and time with a zero mean (Grinsted et al., 2004); it could also be taken as the convolution of the data and the wavelet function (Kovács et al., 2010) for a time series (X_n, $n = 1,. ., N$) with a 'Δt' degree of uniform resolution (Eq. 1):

$$W_n^X(s) = \sqrt{\frac{\Delta t}{S}} \sum_{n'=1}^{N} X_{n'} \Psi_0 \left[(n' - n)\frac{\Delta t}{S} \right] \quad (1)$$

Here N stands for the length of the time series, ψ_0 the wavelet function and s the scale. In the present case to generate daughter wavelets the Morlet mother wavelet (Morlet et al., 1982) was used as the source function.

Wavelet spectrum analysis provides the basis for wavelet transform coherence, which is able to indicate the common power of two variables, being in this way similar to a correlation coefficient, but localized in the frequency-time space (Grinsted et al., 2004). While wavelet spectrum analysis takes into account one variable in 3D (period, power and its localization in the time-frequency space), wavelet transform coherence does the same but for two variables (in this case, one dependent and one independent) in 4D, because the phase differences, which represent the temporal lags, are included as well.

In the study only the positive signals significant ($\alpha = 0.01$) against a thousand first-order auto regressive AR(1), surrogate time series were considered; for details see Roesch and Schmidbauer (2014). It should be noted that, since the wavelet functions at each scale are normalized, the wavelet transforms of the results are comparable even to other time series (Torrence and Compo, 1998). Three main characteristics of the wavelet transform coherence were used:

(i) the presence of the coherent periods in time, which meant that the significant periodic behavior – coherence – at a certain frequency was transformed into percentages, with taking the presence of the coherence/period throughout the whole investigated time as 100% as in previous studies (Hatvani, 2014; Kovács et al., 2010),
(ii) the maximum global–wavelet power, which is the average cross-wavelet power in the frequency domain (averages over time (Roesch and Schmidbauer, 2014)), and
(iii) the phase differences between the pairs of water quality parameters and meteorological parameters which show which series is the leading one in this relationship (Fig. A1).

2.3. Software used

For the calculations, R statistical environment was used (R Core Team, 2016): the wavelet spectrum analysis was performed with the `analyze.wavelet` function, while the wavelet transform coherence results were generated with the `analyze.coherency` function of the `WaveletComp` package (Roesch and Schmidbauer, 2014).

3. Results

3.1. Overview of the system

The varying concentrations of the examined water quality parameters indicate the presence of distinct borders between the different habitats/ecosystems. The River Zala brings a fair amount of nutrients (P and N) to the system through the US site, where about half of the TP is SRP, and where TN mostly consists of NO_3-N (Table 1). In the eutrophic pond these nutrients (SRP; NO_3-N) are mostly bound in algae, which in turn form most of the TSS (Pomogyi, 1996; Fig. A2). Thus, the level of TSS does not significantly decrease compared to that of the River Zala (US), due to the change in its composition from inorganic to organic. In Phase II (WD and DS), however, the amount of N drops to ~50% and TSS to 20% of the concentrations seen in the eutrophic pond, while P retention in Phase II is clearly low (Table 1). It is known that the level of particulate N increases up to the outflow of the eutrophic pond (site AD) then decreases in the wetland (WD); organic matter is decomposed and filtered out by the macrophyte cover (Fig. 2). Dissolved organic nitrogen (DON) shows values similar to that of particulate nitrogen (PN) up to the outflow of the eutrophic pond and accounts for half of TN; it follows the increase of algae biomass (in this case approximated by TSS). At the WD site, DON slightly decreases, but not to the same degree as the nitrate-nitrogen. Therefore, at the downstream outlet of the wetland to Lake Balaton (DS; Fig. 1) N is in dissolved state, but it is not nitrate-nitrogen, rather DON.

The highest degrees of variability (CV > 100%) are reported for TSS in the River Zala (US), and SRP and NO_3-N in the eutrophic pond (AD), and again TSS and NO_3-N in Phase II (WD and DS; Table 1).

Table 1
Descriptive statistics of the water quality parameters at the different sampling locations, where M denotes the mean, SD the standard deviation R the range in mg l^{-1}, and CV the coefficient of variation in% (1994–2010). The number of measurements was equally 6209 for each site and variable.

	Sampling location / water quality parameter	SRP	TP	NO_3-N	TN	TSS
M	US	0.10	0.19	2.01	3.20	33.74
	AD	0.02	0.17	0.42	2.84	24.05
	WD	0.12	0.17	0.22	1.62	3.49
	DS	0.10	0.16	0.27	1.73	5.44
R	US	0.80	3.14	7.83	10.84	3157.00
	AD	0.49	0.83	4.00	12.06	170.00
	WD	0.56	1.07	2.88	12.34	77.00
	DS	0.50	0.86	3.45	8.32	117.00
SD	US	0.06	0.13	0.69	0.97	92.71
	AD	0.03	0.12	0.60	1.43	16.30
	WD	0.10	0.12	0.34	0.60	3.94
	DS	0.08	0.11	0.35	0.62	6.25
CV	US	0.59	0.71	0.34	0.30	2.75
	AD	1.73	0.72	1.44	0.51	0.68
	WD	0.81	0.70	1.53	0.37	1.13
	DS	0.83	0.68	1.31	0.36	1.15

3.2. Periodic behavior of the meteorological parameters

As expected, wavelet spectrum analysis indicated a strong and significant annual periodicity throughout the whole investigated period for all (e.g. Fig. 3a) but one of the independent variables. The exception is precipitation (Fig. 3b). In the power spectrum density graph of precipitation major gaps were observed in its annual periodicity, e.g. between ~2000 and ~2002 (Fig. 3b). In addition, it indicated the weakest global wavelet power in the one-year period band (Table 2). Thus, due to its more intermittent and weak seasonality, it was omitted from the wavelet transform coherence analyses to avoid misleading and unstable results. Regarding the other independent variables, the global wavelet power was highest for T and GR, while the second weakest was for CC. In the case of Q, a clear continuous increase (~34%) can be observed downstream from US to DS (Table 2).

3.3. Common presence of the annual period and maximum global power

As the main step, pairs were set up using the water quality parameters and the independent variables and their coherence was examined using wavelet transform coherence. Results showed that most of the corresponding water quality parameters and the independent variables pairs have a significant common annual periodicity over the entire studied time interval. In the frequency bands other than those corresponding to the annual period, the global wavelet powers of the coherences were always noticeably weak and/or insignificant ($\alpha = 0.01$; as an example, see later Fig. 4a).

This coherence in annual periodicity was most powerful between the P forms and the independent variables (especially GR and T; Table 3). At the US site SRP, and in the eutrophic pond TP, gave a higher global wavelet power at the one-year period band. These powers reached their maxima after the year 2000 (see later Fig. 4). The coherence of P forms with streamflow was the weakest at US and in the AD area, while it was the highest and of the same magnitude in the two sampling locations (WD and DS) of Phase II of the KBWPS. It should be noted that, in general, TP displayed the strongest coherences in the system (avg. global power = 0.70).

Regarding the nitrogen forms, the global wavelet power of TN was of the same magnitude at US and in Phase II (WD and DS); it was strongest at AD with GR and T. In the meanwhile, for NO_3-N, the picture was somewhat similar to that of TN, but more balanced. However, coherence was still highest at AD.

In the case of TSS in general, weak coherences were observed in the system, avg. global power = 0.26 except at AD (Table 3). Its coherence with e.g. CC at US and in Phase II (WD and DS) was < 0.08, making it hard to draw solid conclusions. The highest degrees of coherence were to be seen at AD, where the coherence of the WQPs with CC and Q increased as well. TSS here had nearly as high a degree of coherence with GR and T as did the TP (Table 3). The weakest coherences in general for TSS were seen at WD (avg. power = 0.07).

From the independent variables side, the weakest coherences were observed between the WQPs and CC, and, secondly, with Q. On average, the global wavelet powers were the lowest at US (0.4) and highest at site AD (0.57), while they were of the same magnitude in Phase II (0.51 and 0.52 for WD & DS respectively).

3.3.1. Absence of coherence between the WQPs and the meteorological parameters

Overall, in the whole KBWPS there were 16 occurrences when coherence over an annual scale between the WQPs and the independent variables was interrupted. The absence of annual coherence was only considered if its length was longer than one year, i.e. ~6% of the total investigated time (Table 4). From the perspective of independent variables, these cases were mostly associated with Q (in 12 out of the 17 pairs). Moreover, the highest portion of absence in coherence was usually related to streamflow (~50% of the absence between Q & SRP at AD and

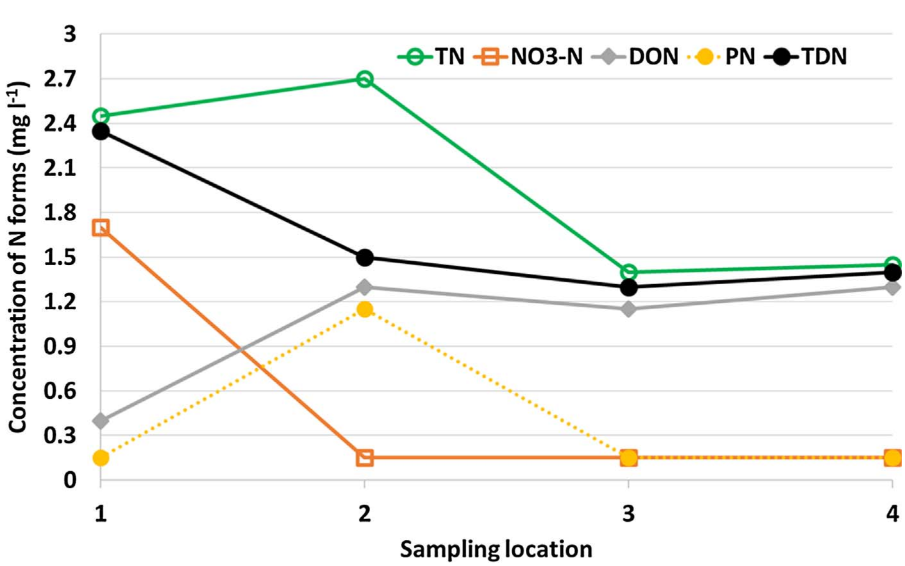


Fig. 2. Average annual (2011) concentrations of TN: total nitrogen; NO_3-N: nitrate-nitrogen; DON: dissolved organic nitrogen; PN: particulate N and TDN: total dissolved N; based on data taken from WTWD (2012).

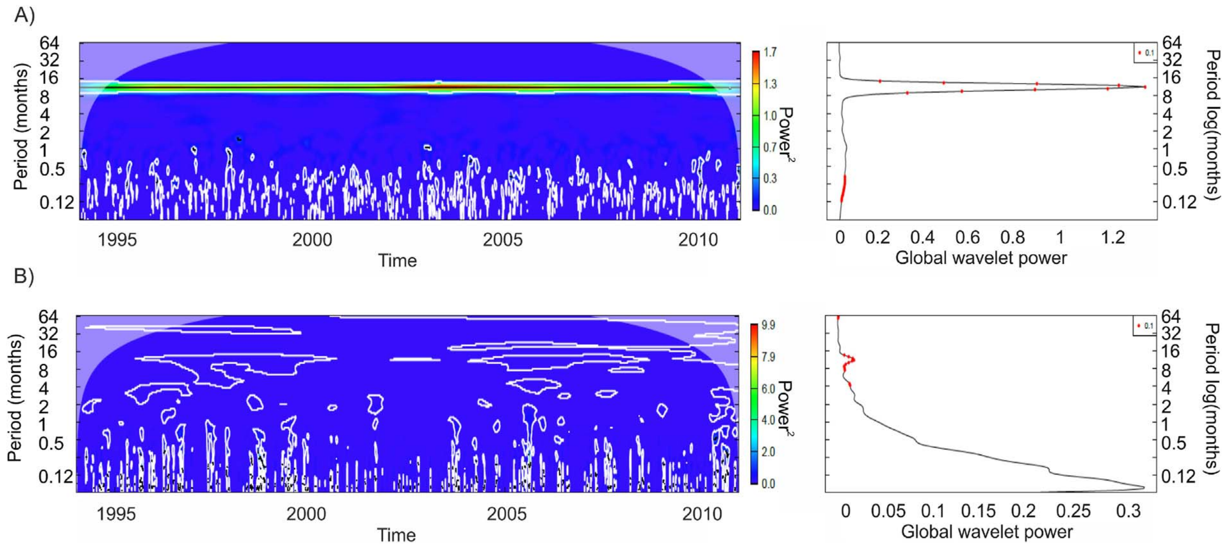


Fig. 3. Power spectrum density (left panels) and time-averaged wavelet power (right panel) graphs indicating the presence of annual periodicity in (a) the temperature and (b) precipitation time series at the US sampling site location for 1994–2010. The white contours in the left panels and the red dots in the right ones show the 90% confidence levels calculated against a thousand AR (1) surrogates. It should be noted that wavelet spectrum analysis coherence and wavelet transform coherence produce edge artifacts, since the wavelet is not completely localized in time, thus the introduction of a cone of influence (COI; dimmed area on the left panels) is suggested, in which edge effects cannot be ignored (Torrence and Compo, 1998). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Global wavelet power of the independent variables at the one-year period for the different sampling site locations (1994–2010; for streamflow (Q) at the WD sampling location 1995–2010).

Water quality parameters / Sampling locations	US	AD	WD	DS
CC	0.18	0.18	0.19	0.18
GR	1.25	1.25	1.25	1.25
Precipitation	0.03	0.03	0.03	0.03
T	1.35	1.40	1.40	1.35
Q	0.23	0.25	0.33	0.35

Q & TSS at US, WD, DS; Table 4). From the perspective of WQPs these episodes of absence in annual coherence were mostly related to SRP and TSS at AD and WD respectively. With regard to the spatial aspect, the average absence decreased in the eutrophic pond and the wetland with respect to the River Zala, after which it increased again at DS (Table 4).

3.4. Phase differences

From the phase differences on the power spectrum density graphs, it is clear that it was mostly the independent variables that were leading the WQPs (e.g. later in Figs. 4, 5 and 7). The phosphorus forms, for example, were mostly in antiphase with CC and Q and in phase with GR and T in the whole system, just as TSS at US and at site AD (Table 5). It was interesting to observe that while T was leading certain WPQs by 1–2 months (e.g. TP at AD; Fig. 4a), GR was leading these by 2–3 months (Fig. 4b). The only habitat where the phase difference of SRP and the independent variables was changing/inconclusive was in the eutrophic pond (AD). TSS in Phase II seems to tend towards keeping the pattern indicated upstream, but its phase differences become changing and inconclusive. It should be noted, that its powers were the lowest here in the whole KBWPS (Table 3).

As for the N forms, NO_3-N, displayed a pattern opposite to that of the P forms (except for SRP and GR at AD). It is in antiphase with T and GR and in-phase with Q, while TN is mostly inconclusive, especially in Phase II

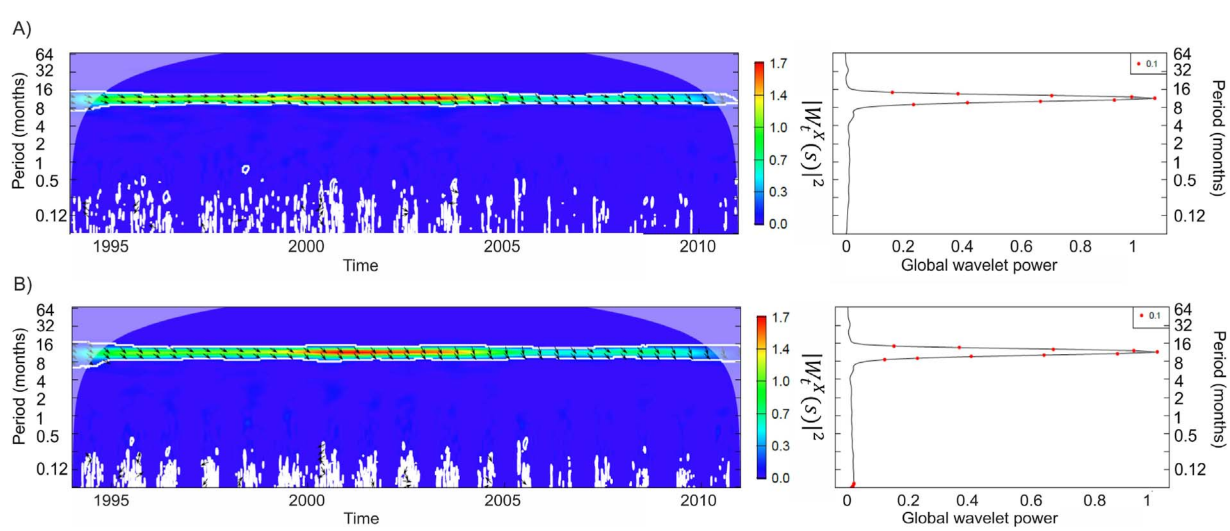


Fig. 4. Time–frequency coherency images (left panel) and time-averaged cross-wavelet power (right panel) of (a) total phosphorus and temperature and (b) global radiation at the AD site. The white contours in the left panels and the red dots in the right ones show the 90% confidence levels calculated against a thousand AR(1) surrogates. The black arrows indicate the phase-angle difference of the parameter pairs. For further details see Roesch and Schmidbauer (2014). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Global wavelet powers of the WQPs and the independent variables (IVs) for 1994–2010. In the case of streamflow (Q), at the WD sampling location, for 1995–2010. The darker red shades indicate higher powers, the darker blue shades smaller ones (For interpretation of the references to colour in this table caption legend, the reader is referred to the web version of this article.).

		Sampling location			
WQP	IVs	US	AD	WD	DS
SRP	CC	0.35	0.15	0.40	0.40
	GR	0.95	0.40	1.10	1.10
	T	1.00	0.48	1.20	1.20
	Q	0.39	0.19	0.53	0.53
TP	CC	0.23	0.38	0.40	0.40
	GR	0.60	1.10	1.10	1.15
	T	0.60	1.10	1.20	1.20
	Q	0.25	0.47	0.54	0.55
NO_3-N	CC	0.25	0.32	0.30	0.28
	GR	0.67	0.80	0.72	0.78
	T	0.68	0.90	0.75	0.80
	Q	0.27	0.39	0.38	0.43
TN	CC	0.15	0.26	0.15	0.14
	GR	0.39	0.78	0.39	0.40
	T	0.42	0.79	0.42	0.40
	Q	0.15	0.29	0.27	0.19
TSS	CC	0.08	0.34	0.04	0.07
	GR	0.25	0.95	0.10	0.18
	T	0.25	0.95	0.10	0.18
	Q	0.10	0.40	0.05	0.10

Table 4
Percentage of the absence of annual coherence for those WQP & independent variable (IV) pairs where the absence was longer than one year (≥6%) of the total time (reference period: 1994–2010; for Q at WD 1995–2010).

		Sampling location			
WQP	IVs	US	AD	WD	DS
SRP	CC		13%		
	GR		11%		
	T		25%		
	Q		56%	7%	
TP	Q		7%		
NO_3-N	Q		20%	11%	
TN	Q	40%	9%	14%	29%
TSS	Q	12%		52%	46%
	T			18%	
Average absence		26%	20%	21%	37%

(Table 5; e.g. Fig. 5a). However, in the River Zala, TN indicates a quasi-persistent antiphase pattern with T, while with GR it was rather hectic (Table 5). It nevertheless showed a quasi-persistent in-phase relationship with T and GR at AD (Fig. 5b). This implies that the N forms besides NO_3-N, organic and particulate, are in-phase with T and GR (Fig. 2).

4. Discussion

4.1. Overview of the coherences

The annual coherence between the water quality- and climatic variables of a river, eutrophic pond and wetland was directly compared. Since there is no transitional area (ecotone) between the ecosystems, the differences in annual coherence clearly represent the distinct habitats, and are as much as possible. In the whole system, the most important factors driving the coherences between the water quality parameters and independent variables were global radiation/temperature, the setting of the different habitats, and the nutrient loads arriving through the River Zala.

In the River Zala (US), annual coherence was strongest between the P forms, nitrate, and T & GR, while absence of coherence was most characteristic of TN and TSS with Q. This can be explained by the general characteristics of small, shallow rivers like the River Zala, with an average depth of 1.4 m at mean water level (GDWM, 2016). The shading effect of riparian vegetation is a key factor in both the heat budget and nutrient cycles of river sections (Allan and Castillo, 2007; Wetzel, 2001). The upper section of the River Zala traverses a forested area, its riparian vegetation shades the water, and the dense canopy prevents excessive warming. The lower section of the river, which is represented by sampling location US, is however much less shaded, since it flows through arable land with scarce riparian vegetation (GDWM, 2016) and with high exposure to heat and radiation, causing the strong coherences with the primary meteorological parameters. Moreover, the fact that nitrate had the weakest antiphase relationship with T and GR in the River Zala can be explained by the generally lower rate and less pronounced seasonal variation of denitrification in rivers compared to lakes (Piña-Ochoa and Álvarez-Cobelas, 2006).

The eutrophic pond (AD) was even more exposed to the effects of air temperature and radiation than the River Zala. The water here slows down, the residence time increases and the pond is slightly shallower than the River Zala (average depth 1.1 m (Tátrai et al., 2000)). With regard to phosphorus, its main processes can be delineated by the Vollenweider model, which describes the relationship of the trophic state of the system based on phosphorus loads and mean depth/retention time (Reynolds, 1992; Vollenweider and Kerekes, 1982). It thus provides ideal conditions for algae to reproduce and consume the SRP in the water (Hatvani et al., 2014) arriving via the River Zala. This is the reason for the lowest SRP values in the whole system (avg. = 0.02 mg l^{-1}; Table 1) being found in the eutrophic pond. In the meanwhile, an opposite process is also present here: with the increase of temperature, the internal P loads of the eutrophic pond increase as well, P is released from the sediment (Istvánovics et al., 2004), especially in drier and warmer years (Chambers and Odum, 1990). This should account for the high degree of coherence between TP (including bounded P in algae cells: “algae-P”) and T & GR in the particularly warm and dry years after 2000 (Fig. 4). These previously discussed processes acting simultaneously (peaking at the same time in the growing season) are responsible for the inconclusive phase difference of SRP and independent variables and the decreased power and occasional absence of their annual coherence. Moreover, since TSS consists mostly of algae in the eutrophic pond (Pomogyi, 1996), it comes as no surprise that the power of its coherence with GR and T was as high as that obtaining between TP and GR & T, because TP consists of “algae-P”. The same notion is true for the N forms as well, especially TN. It predominantly represents the algae – the organic N fraction (Fig. 2) – of the eutrophic pond (Wetzel, 2001). At the same time, inorganic N (nitrate and nitrite) decreased in concentration as SRP, where nitrite was already present in small portions. TSS only indicated a

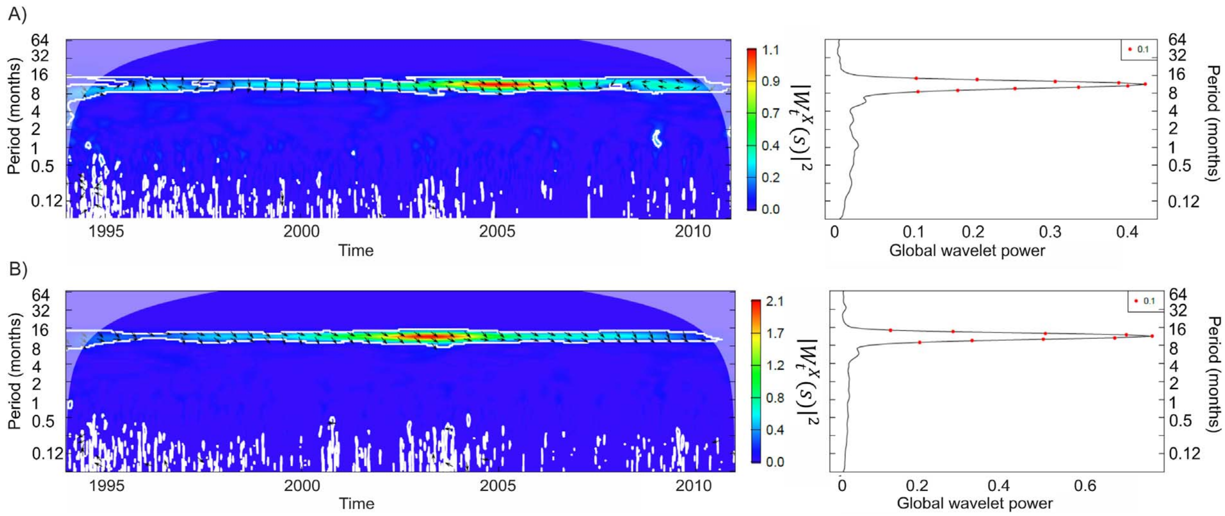


Fig. 5. Time–frequency coherency images (left panel) and the time-averaged cross-wavelet power (right panel) of (a) total nitrogen and temperature at sampling location WD and (b) at sampling location AD. For further details, see the caption to Fig. 4.

strong coherence with the independent variables where it consists mostly of algae; this occurred only in the eutrophic pond.

The waters arriving from eutrophic pond slow down even more and reach the undisturbed- and the disturbed wetland habitat of the KBWPS. Due to the excess loads (see Section 1.1 and Fig. 1), the disturbed "mixed" wetland habitat shows the characteristics of both a classic wetland and a stream. The latter observation manifested itself in the similarity of the disturbed wetland to the River Zala with regard to the global wavelet powers and the absence of annual coherences. In the case of the phase differences, however, the disturbed wetland resembles the classic wetland, indicating that despite the additional inputs both (i.e. the whole of Phase II) are decomposition dominated (Istvánovics et al., 1997), with much lower P retention capacity than the eutrophic pond (Somlyódy, 1998). TN here consists of both organic and inorganic forms, mainly characteristic of processes such as phase changes. Thus, meteorological factors are unlikely to drive TN concentrations. Moreover it has been documented that shading is a factor in dampening the capacity of a wetland to indicate seasonal changes (Kovács et al., 2010) and in controlling the biological processes. It is suspected that the lowest global wavelet power of TN and TSS and the significant gaps in their annual coherence with the independent variables are because of the previously mentioned phenomena. The coherence with the independent variables and the concentration of TSS (Table 1) slightly increases as the additional inputs reach the system. On the one hand, the gaps in annual coherence of TSS and the independent variables were present in the undisturbed wetland because of the mostly low concentrations of TSS (Table 1; Fig. A2) as in macrophyte dominated constructed wetlands (Dunne et al., 2012). While, on the other hand, the gaps between TSS and Q at the output of the system were present due to the unbalanced additional inputs (e.g. Fig. A2) from natural streams, constructed canals and fish ponds (drained three times a year, but irregularly) to Phase II of the KBWPS.

In general, the average percentage of absences in coherency between the water quality parameters and the independent variables decreases as the waters' residence time increases from the River Zala, up to the undisturbed wetland (Section 3.3.1; Table 4). Then, with the additional 40% temporarily irregular input of streamflow downstream of WD, the average percentage of absence increases to values higher than those witnessed in the river. Besides the increased residence time, in the algae dominated eutrophic pond, the cyclic planktonic eutrophication (Wetzel, 2001) played a major role in increasing the ecosystem's capability to follow/indicate meteorological seasonality. A similar pattern was observed by Kovács et al. (2010) in their assessment of annual periodicity using wavelet spectrum analysis on a wider set of weekly sampled parameters for a shorter period (1993–2007). Although in their study the undisturbed wetland showed a higher percentage of absence of annual periodicity (59.1%) than the eutrophic pond (40.9%), still, the disturbed wetland did display a higher absence in annual periodicity (68.2%) than the river (63.6%), as in the present case. The reason for the difference between the obtained absence in periodicity lies not only

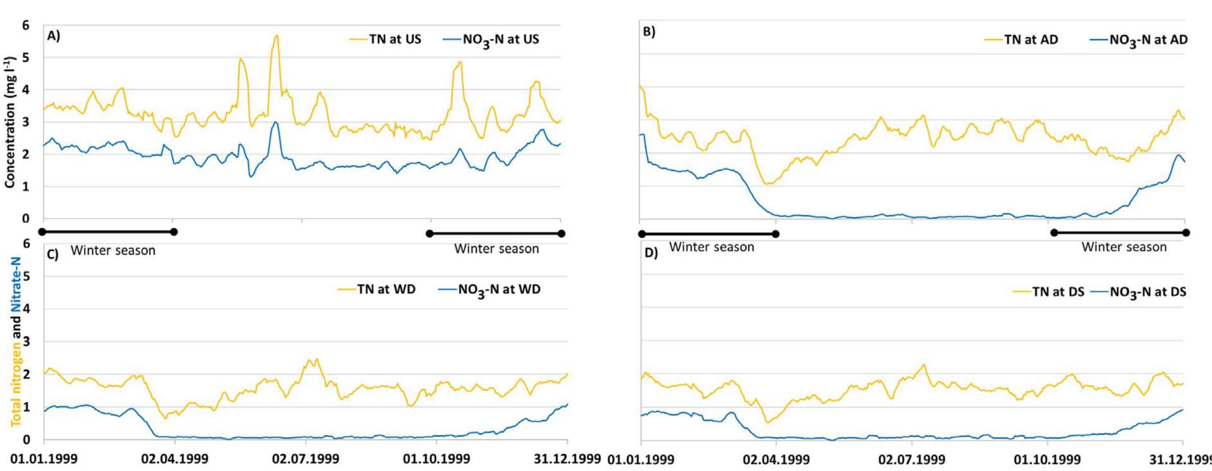


Fig. 6. Centered 7 day moving average of (a) total nitrogen and nitrate-nitrogen in the River Zala, (b) the eutrophic pond, (c) the un-disturbed wetland and (d) the disturbed wetland in 1999. The black lines indicate the winter seasons.

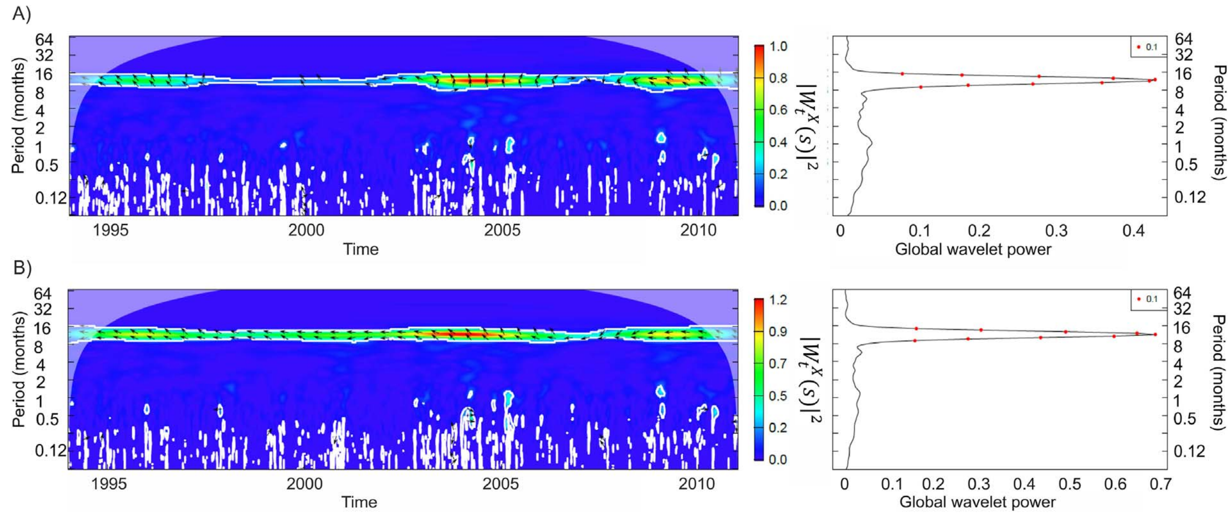


Fig. 7. Time–frequency coherency images (left panel) and time-averaged cross-wavelet power (right panel) of (a) total nitrogen and (b) nitrate-nitrogen with temperature in the River Zala (US). For further details, see the caption to Fig. 4.

Table 5
Phase differences of the WQPs and the independent variables (IVs) for 1994–2010. In the case of Q at sampling location WD, this is for 1995–2010. '−' stands for an antiphase, '+' for an in-phase and IC for an inconclusive/changing phase relationship between the WQPs and the independent variables.

		Sampling location			
WQP	IVs	US	AD	WD	DS
SRP	CC	−	−	−	−
	GR	+	IC	+	+
	T	+	+	+	+
	Q	−	−	−	−
TP	CC	−	−	−	−
	GR	+	+	+	+
	T	+	+	+	+
	Q	−	−	−	−
NO_3-N	CC	+	+	+	+
	GR	−	−	−	−
	T	−	−	−	−
	Q	+	+	+	+
TN	CC	IC	−	IC	IC
	GR	IC	+	IC	IC
	T	−	+	IC	IC
	Q	IC	−	IC	IC
TSS	CC	−	−	IC	IC
	GR	+	+	IC	IC
	T	+	+	IC	IC
	Q	−	−	IC	IC

in the different time interval and applied methodology, but in the fact that the present study focused solely on the nutrient forms and the closely related TSS. The observation that the irregular excess loads arriving to the disturbed wetland corrupt its capability to indicate the seasonal changes emphasizes wetlands' exposure to anthropogenic activity (Brinson and Malvárez, 2002). This vulnerability becomes even more pronounced with climate change (Finlayson, 2016).

4.2. Phase differences

4.2.1. Inconclusive phase differences of P forms and TSS

The pattern of the phase differences concurs with the previously discussed observations; nevertheless, it does provide excess information on the functioning of the system by describing the possible temporal shift between the common annual coherence of the water quality parameters and the independent variables. In the eutrophic pond, TSS for example behaves similarly to TP, being in-phase lead by T (by 1–2 months) and by GR (by 2–3 months), indicating that TSS is composed mostly of algae (Fig. 4), which corresponds to the delay between the weekly average maxima of GR and T. Unsurprisingly, the delay between the two meteorological variables was 7 weeks in the investigated time period, with the GR maxima occurring in the 24th week, i.e. mid-June. By mirroring this meteorologically forced relationship, the capability of the methodology (phase differences) to follow fine changes even under the annual scale is underlined. As for TP, in the eutrophic pond it is most likely to occur in particulate form because of the algae, while in Phase II its wavelet transform coherence results resemble that of SRP, since it is dissolves in the water.

The inconclusive/confusing phase differences between SRP and the independent variables in the eutrophic pond can be explained by the changes in the concentration of P forms through the year, where SRP displayed almost no increase in summer (Fig. A3) due to the continuous algal uptake. Moreover, these inconclusive/confusing phase differences of SRP and T & GR occur for the most part after the year 2000, as was the case of TSS in Phase II. This was a well-documented dry period in the region (Padisák et al., 2006b). In these years, although external nutrient loads decreased, the internal loads acted in the opposite way (Hatvani et al., 2014) due to the higher T and GR. These counter-processes caused e.g. the phase differences of SRP and T to become meaningless, since according to the arrows (Fig. A4), around 2004 T should have been leading SRP by almost 6 months. In the case of TSS, the inconclusive phase differences in Phase II are presumably caused by the generally low concentrations near the level of detection (Fig. A2) and the hectic inputs from the canals.

4.2.2. Inconclusive phase differences of N forms

Upstream, in the River Zala, NO_3-N dominated (Table 1; Figs. 2, 6a) the N forms, with slightly lower concentrations in summer, mostly because the higher exposure of the river section to radiation increases biological activity, thus denitrification (Mulholland et al., 2008). It should be noted, however, that the dissolved organic fraction of TN (Fig. 2) is able to modify the phase differences of TN, and this was especially so in the dry years around 2000 (Fig. 7a). It happened to such an extent that TN was not able to display a pattern (decrease with T and GR in the summer) as clear as in the case of nitrate (Fig. 7b).

In the algae dominated eutrophic pond TN changed its phase with reference to the River Zala, and displayed a clear in phase pattern with T and GR. This occurred because, in the eutrophic pond, as GR and T increase in the growing season, the inorganic N uptake of algae also increases proportionately (Reay et al., 1999). This process decreases the nitrate concentrations (Fig. 6b), thus leaving the TN loads at a similar level as the input from the river (Table 1; Figs. 2 and 6b).

Then the water arrives to the macrophyte covered wetland dominated habitat, where decomposition processes are prevailing (Kovács et al., 2010; Wetzel, 2001), especially in the growing season. Because of the decomposition of algae, oxygen availability is low (Istvanovics, 2002), thus, temperature becomes the most important factor in organic matter loss (Brinson, 1981). If the waters of the River Zala were to enter the wetland directly, probably all N forms would show an opposite/antiphase relationship with T and GR, i.e. lower values in the growing season and higher in winter. This is indeed the case for nitrate (Table 5; Fig. 6c,d), but not for TN, the levels of which do not drop in parallel to this. However, PN is retained by wetlands (Romero et al., 1999) thus decreasing the TN output in the KBWPS accordingly. In summer organic N is continuously resupplied from the decomposition of algae. Therefore, despite the seasonal increase of denitrification (Seitzinger, 1988) and the N uptake of the macrophyte cover (Dvořáková Březinová and Vymazal, 2016) with water temperature, these opposite-tending processes disrupt the periodic characteristic of TN in the wetland area. Nevertheless, a net decrease in the output of TN from the KBWPS is observed due to the previously discussed processes (Table 1); it is just not observable in the seasonal cycle.

5. Conclusions

The water quality variables of a cascade-like engineered ecosystem consisting of a shallow river, a eutrophic pond, and an undisturbed/disturbed macrophyte covered wetland were assessed to track the capacity of the system to indicate meteorological seasonality. In particular, the annual coherence of the water quality parameters and meteorological parameters (including streamflow) indicated the explicit differences in the functioning of the different habitats of the assessed system and these were shown to be in concurrence with previously documented knowledge. It was also pointed out that the eutrophic pond is more capable of mirroring meteorological changes. In the meanwhile, continuous upstream- (from the eutrophic pond) and temporarily irregular additional nutrient inputs (from the southern watershed) tend to counteract the characteristic processes of the wetland (including macrophyte shading). Taken together, these decrease its capacity to indicate seasonality, as seen in the pond upstream. Moreover, it was found that in this particular setting, the wetland is less suitable/unstable in terms of nitrogen retention, and can only decrease the incoming waters' phosphorus concentrations to a small degree, most probably due to the excess- and the high algae loads.

With the successful application of wavelet transform coherence to the "black-box" cascade, where the boxes represent different ecosystems without any transition areas (ecotone) in between them, a promising example is set for the wider application of the method in limnology. The present paper provides a more precise overall picture on the previously discussed behavior of the cascade system, which was designed to restrain the nutrients brought by the River Zala responsible for a fair part of Lake Balaton's eutrophication.

Acknowledgements

The authors would like to thank Paul Thatcher for his work on our English version and say thanks for the stimulating discussions with Péter Tanos from the Szent István University. The work was supported by the MTA "Lendület" program (LP2012-27/2012) and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. This is contribution No. 48 of 2ka Palæoclimatology Research Group.

Appendix

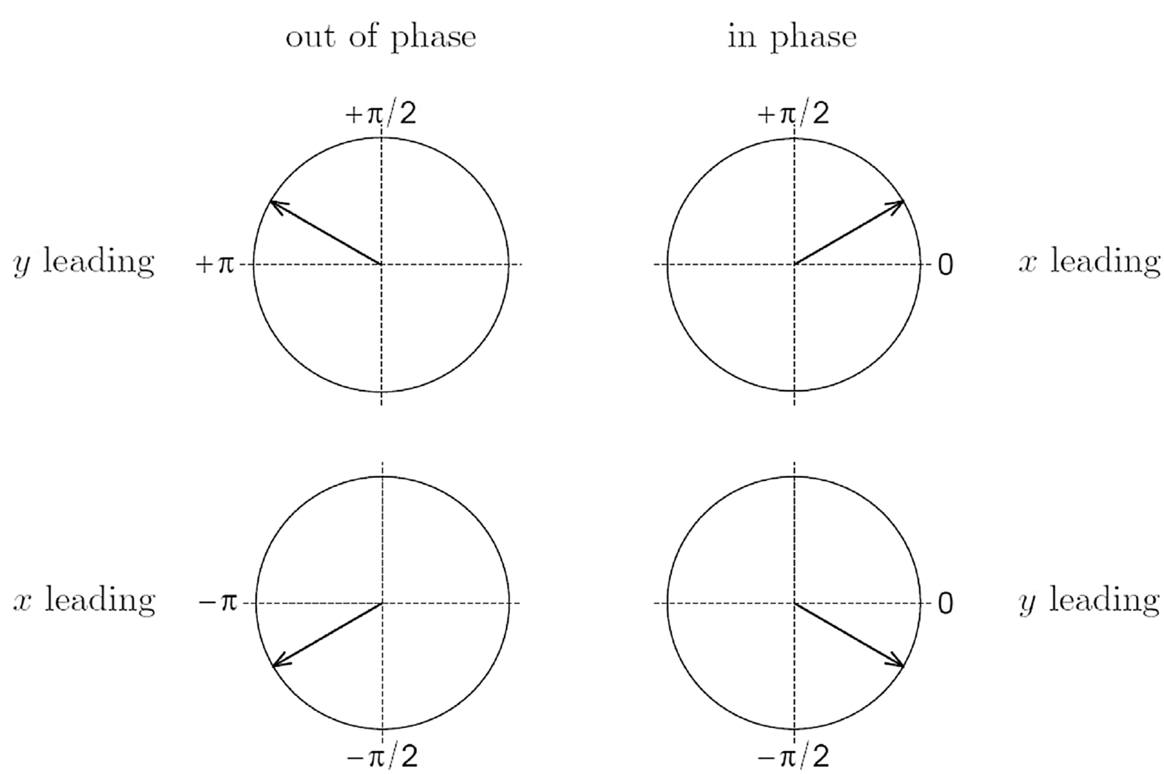


Fig. A1. The full set of possible phase-differences and their interpretation taken from Roesch and Schmidbauer (2014) based on Conraria and Soares (2011), where the phase differences are shown as arrows in the image plot of cross-wavelet power. In the present study the water quality parameter was always the first (x) while the meteorological parameters were the second (y) components of the calculation. In a practical sense for an annual period, the upper left figure would indicate that the meteorological parameter is leading the water quality one in antiphase and with about 2 months; upper right: water quality parameter leading the meteorological one with 2 months in-phase; lower right: water quality parameter antiphase leading the meteorological one with 2 months and lower left: meteorological parameter leading the water quality one with about 2 months in-phase.

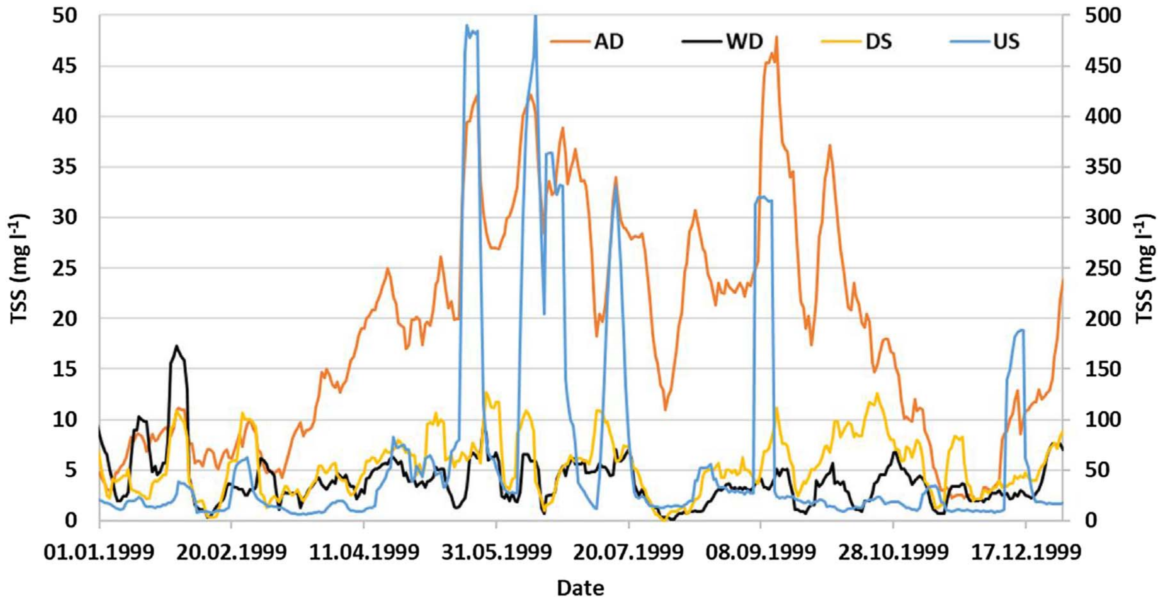


Fig. A2. Centered 7 day moving average of the concentration of total suspended solids in the different habitats of the KBWPS for 1999.

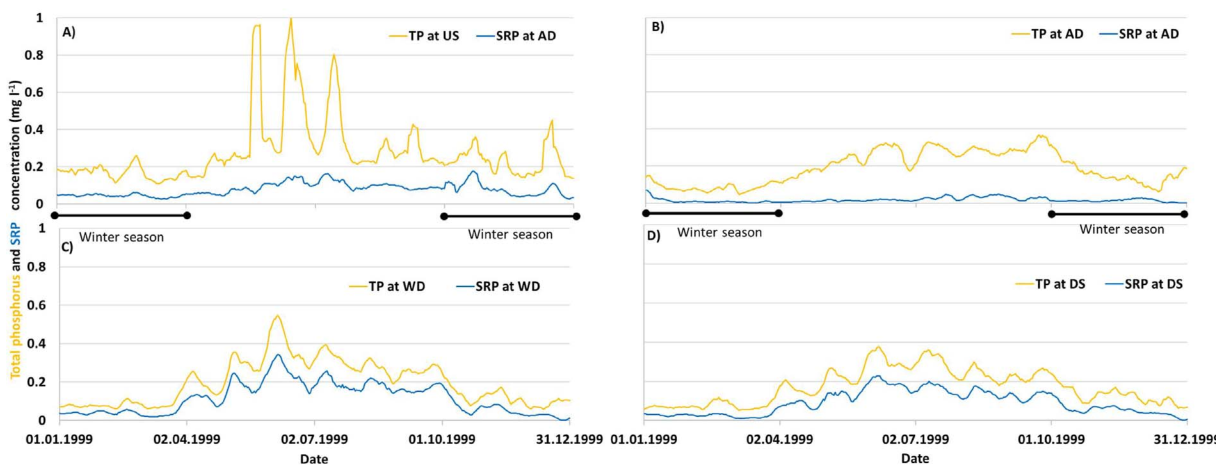


Fig. A3. Centered 7 day moving average of (a) total phosphorus and soluble reactive phosphorus in the River Zala (US), (b) the eutrophic pond (AD), (c) the un-disturbed- (WD) and (d) the disturbed wetland (DS) in 1999. The black horizontal lines indicate the winter seasons.

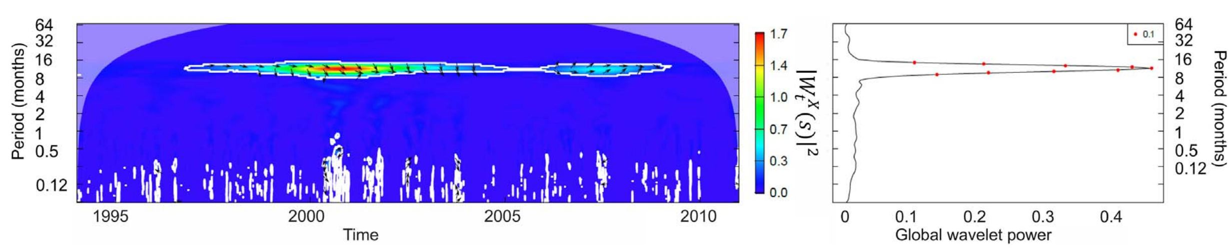


Fig. A4. Time–frequency coherency images (left panel) and time-averaged cross-wavelet power (right panel) of soluble reactive phosphorus and temperature at sampling location AD. For further details, see the caption of Fig. 4.

References

Allan, J.D., Castillo, M.M., 2007. Stream Ecology: Structure and Function of Running Waters. Springer Science & Business Media.
Brinson, M.M., Malvárez, A.I., 2002. Temperate freshwater wetlands: types, status, and threats. Environ. Conserv. 29, 115–133.
Brinson, M.M., 1981. Primary productivity, decomposition and consumer activity in freshwater wetlands. Annu. Rev. Ecol. Syst. 12, 123–161.
Cao, W., Wang, Y., Sun, L., Jiang, J., Zhang, Y., 2016. Removal of nitrogenous compounds from polluted river water by floating constructed wetlands using rice straw and ceramsite as substrates under low temperature conditions. Ecol. Eng. 88, 77–81.
Chambers, R.M., Odum, W.E., 1990. Porewater oxidation, dissolved phosphate and the iron curtain. Biogeochemistry 10, 37–52.
Conraria, L.A., Soares, M.J., 2011. The continuous wavelet transform: a primer. NIPE Work. Pap. 16, 1–43.
Dettinger, M.D., Diaz, H.F., 2000. Global characteristics of stream flow seasonality and variability. J. Hydrometeorol. 1, 289–310.
Dunne, E.J., Coveney, M.F., Marzolf, E.R., Hoge, V.R., Conrow, R., Naleway, R., Lowe, E.F., Battoe, L.E., 2012. Efficacy of a large-scale constructed wetland to remove phosphorus and suspended solids from Lake Apopka, Florida. Ecol. Eng. 42, 90–100.
Dunne, E.J., Coveney, M.F., Hoge, V.R., Conrow, R., Naleway, R., Lowe, E.F., Battoe, L.E., Wang, Y., 2015. Phosphorus removal performance of a large-scale constructed treatment wetland receiving eutrophic lake water. Ecol. Eng. 79, 132–142.
Dvořáková Březinová, T., Vymazal, J., 2016. Distribution of Phosphorus and Nitrogen in Phragmites Aboveground Biomass. In: Vymazal, J. (Ed.), Natural and Constructed

Wetlands. Springer, Cham, pp. 69–76.
EC, 2000. Directive 2000/60/European Commission of the European Parliament and of the Council of 23 October 2000 establishing a framework for community action in the field of water policy. Off. J. Eur. Commun. L 327, 1–72.
Exner-Kittridge, M., Strauss, P., Blöschl, G., Eder, A., Saracevic, E., Zessner, M., 2016. The seasonal dynamics of the stream sources and input flow paths of water and nitrogen of an Austrian headwater agricultural catchment. Sci. Total Environ. 542 (Part A), 935–945.
Finlayson, C.M., 2016. Climate change and wetlands. In: Finlayson, C.M., Everard, M., Irvine, K., McInnes, R.J., Middleton, B.A., van Dam, A.A., Davidson, N.C. (Eds.), The Wetland Book: I: Structure and Function, Management and Methods. Springer, Netherlands, Dordrecht, pp. 1–12.
GDWM, 2016. Water Management Plan 2015- Hungarian Part of the Danube River Basin (in Hungarian). General Directorate of Water Management p. 698.
Grinsted, A., Moore, J.C., Jevrejeva, S., 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. Nonlinear Processes Geophys. 11, 561–566.
Hatvani, I.G., Clement, A., Kovács, J., Kovács, I.S., Korponai, J., 2014. Assessing water-quality data: the relationship between the water quality amelioration of Lake Balaton and the construction of its mitigation wetland. J. Great Lakes Res. 40, 115–125.
Hatvani, I.G., Kovács, J., Márkus, L., Clement, A., Hoffmann, R., Korponai, J., 2015. Assessing the relationship of background factors governing the water quality of an agricultural watershed with changes in catchment property (W-Hungary). J. Hydrol. 521, 460–469.
Hatvani, I.G., Kern, Z., Leél-Őssy, S., Demény, A., 2017. Speleothem stable isotope reference records for East-Central Europe – resampling sedimentary proxy records to get evenly spaced time-series with spectral control. Earth Syst. Sci. Data Discuss. 1–17.
Hatvani, I.G., 2014. Application of State-of-the-Art Geomathematical Methods in Water Protection: – On the Example of the Data Series of the Kis-Balaton Water Protection System, School of Environmental Sciences. Eötvös Loránd University p. 110.
ICPDR, 2015. The Danube River Basin District Management Plan. ICPDR – International Commission for the Protection of the Danube River, Vienna p. 192.
Istvánovics, V., Kovács, A., Voros, L., Herodek, S., Pomogyi, P., 1997. Phosphorus cycling in a large, reconstructed wetland, the lower Kis-Balaton Reservoir (Hungary). Verhandlungen-Internationale Vereinigung für theoretische und angewandte Limnologie 26, 323–329.
Istvánovics, V., Osztoics, A., Honti, M., 2004. Dynamics and ecological significance of daily internal load of phosphorus in shallow Lake Balaton, Hungary. Freshw. Biol. 49, 232–252.
Istvánovics, V., Clement, A., Somlyódy, L., Specziár, A., G.-Tóth, L., Padisák, J., 2007. Updating water quality targets for shallow Lake Balaton (Hungary), recovering from eutrophication. In: Qin, B., Liu, Z., Havens, K. (Eds.), Eutrophication of Shallow Lakes with Special Reference to Lake Taihu, China. Springer, Netherlands, Dordrecht, pp. 305–318.
Istvánovics, V., 1988. Seasonal variation of phosphorus release from the sediments of Shallow Lake Balaton (Hungary). Water Res. 22, 1473–1481.
Istvanovics, V., 2002. Pretention as a function of environmental conditions in a large, reed-dominated wetland. Internationale Vereinigung fur Theoretische und Angewandte Limnologie Verhandlungen 27, 4005–4008.
Kadlec, R.H., 1999. Chemical, physical and biological cycles in treatment wetlands. Water Sci. Technol. 40, 37–44.
Kolander, R., Tylkowski, J., 2008. Hydrochemical seasons in the Lake Gardno catchment on Wolin Island (north-western Poland). Limnol. Rev. 8, 27–34.
Kovács, J., Hatvani, I.G., Korponai, J., Kovács, I.S., 2010. Morlet wavelet and autocorrelation analysis of long-term data series of the Kis-Balaton water protection system (KBWPS). Ecol. Eng. 36, 1469–1477.
Kovács, J., Tanos, P., Várbíró, G., Anda, A., Molnár, S., Hatvani, I.G., 2017. The role of annual periodic behavior of water quality parameters in primary production: chlorophyll-a estimation. Ecol. Indic. 78C, 311–321.
Lotz, G., 1988. The Kis-Balaton water protection system [in Hungarian]. Hidrologiai Tájékoztató 10, 20–22.
Mátyás, K., Oldal, I., Korponai, J., Tátrai, I., Paulovits, G., 2003. Indirect effect of different fish communities on nutrient chlorophyll relationship in shallow hypertrophic water quality reservoirs. Hydrobiologia 504, 231–239.
Martín, M., Oliver, N., Hernández-Crespo, C., Gargallo, S., Regidor, M.C., 2013. The use of free water surface constructed wetland to treat the eutrophicated waters of lake L'Albufera de Valencia (Spain). Ecol. Eng. 50, 52–61.
Morlet, J., Arens, G., Fourgeau, E., Giard, D., 1982. Wave propagation and sampling theory; part I, complex signal and scattering in multilayered media. Geophysics 47, 203–221.
Mulholland, P.J., Helton, A.M., Poole, G.C., Hall, R.O., Hamilton, S.K., Peterson, B.J., Tank, J.L., Ashkenas, L.R., Cooper, L.W., Dahm, C.N., Dodds, W.K., Findlay, S.E.G., Gregory, S.V., Grimm, N.B., Johnson, S.L., McDowell, W.H., Meyer, J.L., Valett, H.M., Webster, J.R., Arango, C.P., Beaulieu, J.J., Bernot, M.J., Burgin, A.J., Crenshaw, C.L., Johnson, L.T., Niederlehner, B.R., O/'Brien, J.M., Potter, J.D., Sheibley, R.W., Sobota, D.J., Thomas, S.M., 2008. Stream denitrification across biomes and its response to anthropogenic nitrate loading. Nature 452, 202–205.
Nalley, D., Adamowski, J., Khalil, B., Biswas, A., 2016. Inter-annual to inter-decadal streamflow variability in Quebec and Ontario in relation to dominant large-scale climate indices. J. Hydrol. 536, 426–446.
Ni, J., Xu, J., Zhang, M., 2016. Constructed wetland planning-based bi-level optimization to balance the watershed ecosystem and economic development: a case study at the Chaohu Lake watershed, China. Ecol. Eng. 97, 106–121.
Padisák, J., Borics, G., Grigorszky, I., Soróczki-Pintér, É., 2006a. Use of phytoplankton assemblages for monitoring ecological status of lakes within the water framework directive: the assemblage index. Hydrobiologia 553, 1–14.
Padisák, J., Molnár, G., Soróczki-Pintér, É., Hajnal, É., George, D.G., 2006b. Four consecutive dry years in Lake Balaton (Hungary): consequences for phytoplankton biomass and composition. Verhandlungen-Internationale Vereinigung für theoretische und angewandte Limnologie 29, 1153–1159.
Padisak, J., 1992. Seasonal succession of phytoplankton in a large shallow lake (Balaton, Hungary) – a dynamic approach to ecological memory, its possible role and mechanisms. J. Ecol. 80, 217–230.
Piña-Ochoa, E., Álvarez-Cobelas, M., 2006. Denitrification in aquatic environments: a cross-system analysis. Biogeochemistry 81, 111–130.
Pomogyi, P., 1996. Second Kis-Balaton Conference: Summary of the Research Result on the KBWPS from Between 1991–1995. University of Pannonia Georgikon Faculty of Agriculture, Keszthely.
R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Reay, D.S., Nedwell, D.B., Priddle, J., Ellis-Evans, J.C., 1999. Temperature dependence of inorganic nitrogen uptake: reduced affinity for nitrate at suboptimal temperatures in both algae and bacteria. Appl. Environ. Microbiol. 65, 2577–2584.
Reynolds, C.S., 1984. Phytoplankton periodicity: the interactions of form, function and environmental variability. Freshw. Biol. 14, 111–142.
Reynolds, C., 1992. Eutrophication and the Management of Planktonic Algae: What Vollenweider Couldn't Tell Us, Eutrophication: Research and Application to Water Supply. Freshwater Biological Association, Ambleside, UK, pp. 4–29.
Roesch, A., Schmidbauer, H., 2014. WaveletComp: A Guided Tour Through the R-package. p. 38.
Romero, J.A., Comín, F.A., García, C., 1999. Restored wetlands as filters to remove nitrogen. Chemosphere 39, 323–332.
Salamalikis, V., Argiriou, A.A., Dotsika, E., 2016. Periodicity analysis of δ18O in precipitation over Central Europe: time–frequency considerations of the isotopic 'temperature' effect. J. Hydrol. 534, 150–163.
Seitzinger, S.P., 1988. Denitrification in freshwater and coastal marine ecosystems: ecological and geochemical significance. Limnol. Oceanogr. 33, 702–724.
Sen, A.K., Kern, Z., 2016. Wavelet analysis of low-frequency variability in oak tree-ring chronologies from east Central Europe. Open Geosci. 8, 478–483.
Somlyódy, L., Herodek, S., Fischer, J., Akadémia, M.T., Analysis, I.I.f.A.S., 1983. Eutrophication of shallow lakes: modeling and management: the lake balaton case study. Proceedings of a Workshop International Institute for Applied Systems Analysis.
Somlyódy, L., 1998. Eutrophication modeling, management and decision making: the Kis-Balaton case. Water Sci. Technol. 37, 165–175.
Spinoni, J., Szalai, S., Szentimrey, T., Lakatos, M., Bihari, Z., Nagy, A., Németh, Á., Kovács, T., Mihic, D., Dacic, M., Petrovic, P., Kržič, A., Hiebl, J., Auer, I., Milkovic, J., Štepánek, P., Zahradnícek, P., Kilar, P., Limanowka, D., Pyrc, R., Cheval, S., Birsan, M.-V., Dumitrescu, A., Deak, G., Matei, M., Antolovic, I., Nejedlík, P., Štastný, P., Kajaba, P., Bochnícek, O., Galo, D., Mikulová, K., Nabyvanets, Y., Skrynyk, O., Krakovska, S., Gnatiuk, N., Tolasz, R., Antofie, T., Vogt, J., 2015. Climate of the Carpathian Region in the period 1961–2010: climatologies and trends of 10 variables. Int. J. Climatol. 35, 1322–1341.
Tátrai, I., Mátyás, K., Korponai, J., Paulovits, G., Pomogyi, P., 2000. The role of the Kis-Balaton Water Protection System in the control of water quality of Lake Balaton. Ecol. Eng. 16, 73–78.
Tanos, P., Kovács, J., Kovács, S., Anda, A., Hatvani, I., 2015. Optimization of the monitoring network on the River Tisza (Central Europe, Hungary) using combined cluster and discriminant analysis, taking seasonality into account. Environ. Monit. Assess. 187, 1–14.
Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. Bull. Am. Meteorol. Soc. 79, 61–78.
Vörösmarty, C.J., Green, P., Salisbury, J., Lammers, R.B., 2000. Global water resources: vulnerability from climate change and population growth. Science 289, 284–288.
Vollenweider, R.A., Kerekes, J., 1982. Eutrophication of Waters. Monitoring, Assessment and Control. OECD Cooperative Programme on Monitoring of Inland Waters (Eutrophication Control). Environment Directorate OECD.
WTWD, West-Transdanubian Water Directorate, 2012. Annual Report on the Results of the Water Qualtiy Analyses of the KBWPS (in Hungarian). p. 84.
Wang, F., Wang, X., Zhao, Y., Yang, Z., 2012. Long-term changes of water level associated with chlorophyll a concentration in lake Baiyangdian, North China. Procedia Environ. Sci. 13, 1227–1237.
Wetzel, R.G., 2001. Limnology. Academic Press, San Diego.
Wong, S.L., Clark, B., Kirby, M., Kosciuw, R.F., 1978. Water temperature fluctuations and seasonal periodicity of cladophora and potamogeton in shallow rivers. J. Fish. Res. Board Can. 35, 866–870.

Environmental Science and Pollution Research (2018) 25:28884–28898
https://doi.org/10.1007/s11356-018-2783-7

RESEARCH ARTICLE

Hotspots and main drivers of fecal pollution in Neusiedler See, a large shallow lake in Central Europe

István G. Hatvani[1] · Alexander K. T. Kirschner[2,3,4] · Andreas H. Farnleitner[3,5] · Péter Tanos[6] · Alois Herzig[7,8]

Received: 24 April 2018 / Accepted: 16 July 2018 / Published online: 13 August 2018

Abstract
To minimize the risk of negative consequences for public health from fecal pollution in lakes, the continuous surveillance of microbiological water quality parameters, alongside other environmental variables, is necessary at defined bathing sites. Such routine surveillance may prove insufficient to elucidate the main drivers of fecal pollution in a complex lake/watershed ecosystem, and it may be that more comprehensive monitoring activities are required. In this study, the aims were to identify the hotspots and main driving factors of fecal pollution in a large shallow Central European lake, the Neusiedler See, and to determine to what degree its current monitoring network can be considered representative spatially. A stochastic and geostatistical analysis of a huge data set of water quality data (~ 164,000 data points, representing a 22-year time-series) of standard fecal indicator bacteria (SFIB), water quality and meteorological variables sampled at 26 sampling sites was conducted. It revealed that the hotspots of fecal pollution are exclusively related to sites with elevated anthropogenic activity. Background pollution from wildlife or diffuse agricultural run-off at more remote sites was comparatively low. The analysis also showed that variability in the incidence of SFIB was driven mainly by meteorological phenomena, above all, temperature, number of sunny hours, and wind (direction and speed). Due to antagonistic effects and temporal undersampling, the influence of precipitation on SFIB variance could not be clearly determined. Geostatistical analysis did reveal that the current spatial sampling density is insufficient to cover SFIB variance over the whole lake, and that the sites are therefore in the most part representative of local phenomena. Suggestions for the future monitoring and managing of fecal pollution are offered. The applied statistical approach may also serve as a model for the study of other such areas, and in general indicate a method for dealing with similarly large and spatiotemporally heterogeneous datasets.

Keywords Large lake · Microbial fecal pollution · Monitoring · Principal component analysis · Stochastic modeling · Variography · Kriging

István G. Hatvani and Alexander K. T. Kirschner contributed equally to this work.

Responsible editor: Philippe Garrigues

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11356-018-2783-7) contains supplementary material, which is available to authorized users.

✉ Alexander K. T. Kirschner
alexander.kirschner@meduniwien.ac.at;
https://www.waterandhealth.ac.at

[1] Institute for Geological and Geochemical Research, Research Centre for Astronomy and Earth Sciences, Hungarian Academy of Sciences (MTA), Budaörsi út 45, Budapest H-1112, Hungary

[2] Institute for Hygiene and Applied Immunology—Water Hygiene, Medical University Vienna, Kinderspitalgasse 15, A-1090 Vienna, Austria

[3] Karl Landsteiner University for Health Sciences, Dr.-Karl-Dorrek-Straße 30, A-3500 Krems, Austria

[4] Interuniversity Cooperation Centre Water & Health, Vienna, Austria

[5] Technische Universität Wien, Research Centre for Water and Health 057-08, Institute for Chemical, Environmental and Bioscience Engineering, Gumpendorferstrasse 1a, A-1060 Vienna, Austria

[6] Department of Mathematics and Informatics, Szent István University, Páter Károly utca 1, Gödöllő H-2100, Hungary

[7] Biological Research Institute Burgenland, A-7142 Illmitz, Austria

[8] Nationalpark Neusiedler See-Seewinkel, A-7143 Apetlon, Austria

Introduction

Lakes in the temperate climate zone are intensively used for recreational purposes. The changes witnessed in the ecological status of European lakes over the second half of the twentieth century severely impacted their usability for water supply, recreation, and tourism (Dokulil 2014; MEA 2005) and their ecosystem services. However, due to the comprehensive measures, e.g., (1) the construction of sewage systems, waste water treatment plants (WWTPs) (see, for example, Hatvani et al. 2014a), (2) the introduction of natural protection zones, e.g., Hatvani et al. (2014b), and (3) the implementation of the EU Water Framework Directive (EU-WFD) (EC 2000) etc., water quality has improved to a marked degree in recent decades. With regard to fecal pollution, specific water quality targets have been implemented in legislation which demands the surveillance of water quality on a regular basis to enable safe recreational water use (EU Bathing Water Directive, EU-BWD, EC 2006). In addition to surveillance, a fundamental understanding of potential sources of pollution and of the factors influencing pollution patterns in lakes is a prerequisite for effective water management. In all European Union countries, the surveillance of microbiological water quality extends over the summer bathing season and encompasses the enumeration of standard fecal indicator bacteria (SFIB) *Escherichia coli* and intestinal enterococci, based on international standard procedures; this is conducted on a relatively small number of occasions at defined time intervals at nominated bathing water sites (EC 2006).

The main aim of the directive has been to force EU member states to minimize the risk of negative consequences for public health from fecal pollution in bathing waters through improved water monitoring and management practices (Bedri et al. 2016). Due to (1) the increased awareness of water quality deterioration in lakes and rivers (Strobl and Robillard 2008), and (2) developments in analytical and monitoring equipment, the number of monitoring sites and water quality parameters assessed (physical, chemical, and biological) has increased in rivers (Chapman et al. 2016) and lakes (Kovács et al. 2012b) as well. This has, in turn, led to increased opportunities to spot and manage fecal pollution in waters used for recreation. However, improper monitoring practices and the under-exploitation of the information contained in the data clearly undermine these aims (Ward et al. 1986). The proper planning of monitoring strategies for rivers and lakes—e.g., Bartram et al. (1996), Chapman et al. (2016), Strobl and Robillard (2008)—and the thorough statistical processing of water quality data from activities related to the EU-WFD and EU BWD is essential if representative information is to be obtained (Hatvani et al. 2014b; Wymer 2007). This is a vital requirement in international guidelines for bathing water management (WHO 2003.)

In Austria, since the effective implementation of wastewater management in the 1970s and 1980s, the chemical and microbiological water quality of lakes has significantly improved (Dokulil 2017). Besides the construction of WWTPs, water management is based on the rerouting of contaminated wastewater away from the lake ecosystem and the discharge of the treated wastewater into flowing waters (rivers). The situation at the Neusiedler See, the western-most brackish lake in Europe and the largest lake in Austria, is, however, different (Magyar et al. 2013). The lake is under the de-jure protection of the Ramsar Convention (Convention on Wetlands 1971) and lies within the territory of national parks in both Austria and Hungary; furthermore, the site gained World Heritage status in 2001 (Dinka et al. 2016). Nonetheless, it receives treated wastewater directly from four WWTPs (one on the Hungarian part of the lake). The outflow of further three WWTPs enters the lake through the waters of the River Wulka (Magyar et al. 2013). Moreover, unrecognized point and non-point sources (agricultural run-off, livestock, and wildlife) may also add to the fecal pollution of the lake. According to the EU-BWD (EC 2006), routine surveillance of bathing water sites in the lake is restricted annually to five samples, the minimum required, collected during the bathing season for four consecutive years to obtain a classification of the bathing water quality. Prior to the implementation of the revised EU-BWD directive into Austrian national legislation (Badegewässerverordnung 2009, BGBl II/349), different microbiological parameters such as fecal coliforms, total coliforms, or streptococci were used by different member countries, while surveillance practices differed considerably between them.

Since routine surveillance does not necessarily allow an understanding of the observed complex patterns of fecal pollution in the lake, a much more detailed monitoring of microbial water quality parameters and their relation to potential influencing factors (e.g., meteorological parameters) is required. Thus, in order to expand the routine surveillance of bathing sites during the bathing season at Neusiedler See, an extended set of parameters for microbiological water quality (SFIB) covering 26 critical sampling sites around the lake over a period of 22 years was examined. Samples were taken not only during the bathing season (June 1st to August 31st) but during the period from March 1st to October 31st, when several recreational activities take place (surfing, sailing, fishing, etc.), in the course of which visitors come in direct contact with the water. Preliminary observations did shed light on a few alarming water quality events and a large data set has been collected over time. Still, no thorough statistical investigation has been conducted to assess the lake-wide patterns and driving factors of SFIB variance in order to guide future monitoring and management practices.

Therefore, stochastic and geostatistical analyses were employed (1) to identify the hotspots of fecal pollution in the Neusiedler See based on SFIB abundance, (2) to elucidate the main driving factors of the fecal pollution, and (3) to assist in the optimization of the current spatiotemporal state of the monitoring system with respect to its spatial sampling frequency. With this comprehensive approach, a better management of the microbiological water quality of Neusiedler See should be possible, and it is hoped that this study will serve as a model for comparable lake ecosystems.

Materials and methods

Study area description

The Neusiedler See (47°42′ N, 16°46′ E) is the largest (full area 309 km^2) shallow (average water depth 1.1 m) lake in Austria, though a quarter of this area actually lies within the territory of Hungary (75 km^2) (Dinka et al. 2016). Its main water inputs are precipitation (average 181×10^6 m^3 $year^{-1}$, 78% for 1967–2012; Kubu et al. 2014) and the two small natural streams, the Wulka (Austria, average discharge 42×10^6 m^3 $year^{-1}$, 14%) and the Rákos (Hungary; average discharge 2.5×10^6 m^3 $year^{-1}$, 1%); the Golser Canal (Austria, average discharge 7×10^6 m^3 $year^{-1}$, 2.5%) also plays a role. Its output is mainly due to evaporation and evapotranspiration (208×10^6 m^3 $year^{-1}$, 90%). About 10% of its shoreline is affected by anthropogenic activity. In periods of high water levels (above 115.8 m.a.s.l.), a channel is activated that allows the run-off of surplus water (26×10^6 m^3 $year^{-1}$, 10%). There is almost constant air movement in the area, and thanks to orographic effects the prevailing wind direction is NW (Józsa et al. 2008). The lake has dried out many times over the centuries. Therefore, water management measures are continuously applied, with a considerable influence on the current form of the reed stands which cover about 56% of the lake (Magyar et al. 2013).

A major driver of the microbiological-hygienic status of the lake is the inflow of treated wastewater (Kirschner et al. 2014). Several indirect and direct wastewater influents from WWTPs into the lake exist around the lake (Table S1). The major indirect influent is the Wulka, with approximately 220,000 population equivalents (PE). The major direct influent is the WWTP at Podersdorf (PE 20,000), which is the only WWTP equipped with UV disinfection. All of these influents pass through the reed stands. However, the width of the reed stands on the west side of the lake is much greater than on the east side. In the west, they extend over 2–3 km, receiving the Wulka, the effluents of Jois (PE 8400) and Balf (PE 4000), while in the east they extend only over 50–100 m, receiving the Golser Canal (PE 40,000) and the effluents of Podersdorf. In the case of all WWTP influents, sewer overflows in conditions of heavy rainfall are considered as potential main pollution events for the lake (Sommer et al. 2018). In addition to the WWTPs, the inflow of polluted water may derive from smaller channels draining agricultural areas (mainly cattle and horse raising), from surface run-off during rainfall events, and from wildlife (mainly birds).

Description of sampling sites

In total, eight official EU bathing sites are situated around the Neusiedler See, seven in Austria, in the municipalities of Mörbisch, Rust, Breitenbrunn, Neusiedl, Weiden, Podersdorf and Illmitz, and one in Hungary, Tavi-Vizitelep (EEA 2018). In this study, only the Austrian sites were considered because for the Hungarian part of the lake, the set of parameters assessed, the sampling strategy, and the analytical methods used differed from the Austrian ones, rendering their joint handling impossible. In addition to the seven Austrian official bathing sites, further 19 sites were selected for the monitoring of microbial fecal pollution, those having sufficient available data for the whole investigation period, 1992–2013 (Table S1; Fig. 1). They included potentially critical sites of microbial fecal pollution and a few reference sites with "background" levels of fecal pollution.

Acquired dataset and preprocessing

Water quality data

The time series of the analyzed water quality parameters, including SFIB data, was acquired from the Biological Research Institute Burgenland at Illmitz. The presence of intestinal enterococci was determined in accordance with ISO 7899–2:2000. From 2003 onwards, *E. coli* concentrations were determined using the Colilert-18 system (IDEXX, Ludwigsburg, Germany), which became an ISO standard procedure in 2012 (ISO 9308–2:2012). Prior to 2003, in the place of *E. coli*, thermotolerant coliforms were recorded, in accordance with EN ISO 9308–1:1990 (Part 1, membrane filtration method). In 2003, both methods were run in tandem. Due to this change in the analytical method, a conversion to Colilert based on the parallel measurements of data dating from before 2003 had to be made; for details see Electronic Supplementary Material (ESM) Section "Conversion of *E.coli* concentration." All SFIB are expressed in CFU (colony forming units, Enterococci) or MPN (most probable number, *E.coli*) per 100 ml and $\log_{10}$ transformed for statistical analysis. The other water quality parameters examined were water temperature (W_{Temp}; °C), pH, electric conductivity (Cond.; mS cm^{-1}), dissolved oxygen (DO), total phosphorus (TP), ammonium-N, nitrate-N, chlorophyll-*a* (all mg l^{-1}), and Secchi depth (cm), all collected at 26 sampling sites (Fig. 1). During winter, no sampling was conducted at many monitoring sites. Thus, for each year from 1992 to 2013, all the data (water quality and SFIB) measured between March 1st and October 31st (at weekly/bi-weekly/monthly intervals, depending on the station and the year) were averaged for each year and included in the analysis. The same procedure was conducted on the water level data (cm above Adriatic Sea level) recorded daily at six sites across the lake (Apetlon, Breitenbrunn, Illmitz, Mörbisch, Podersdorf, and Rust), along with the daily

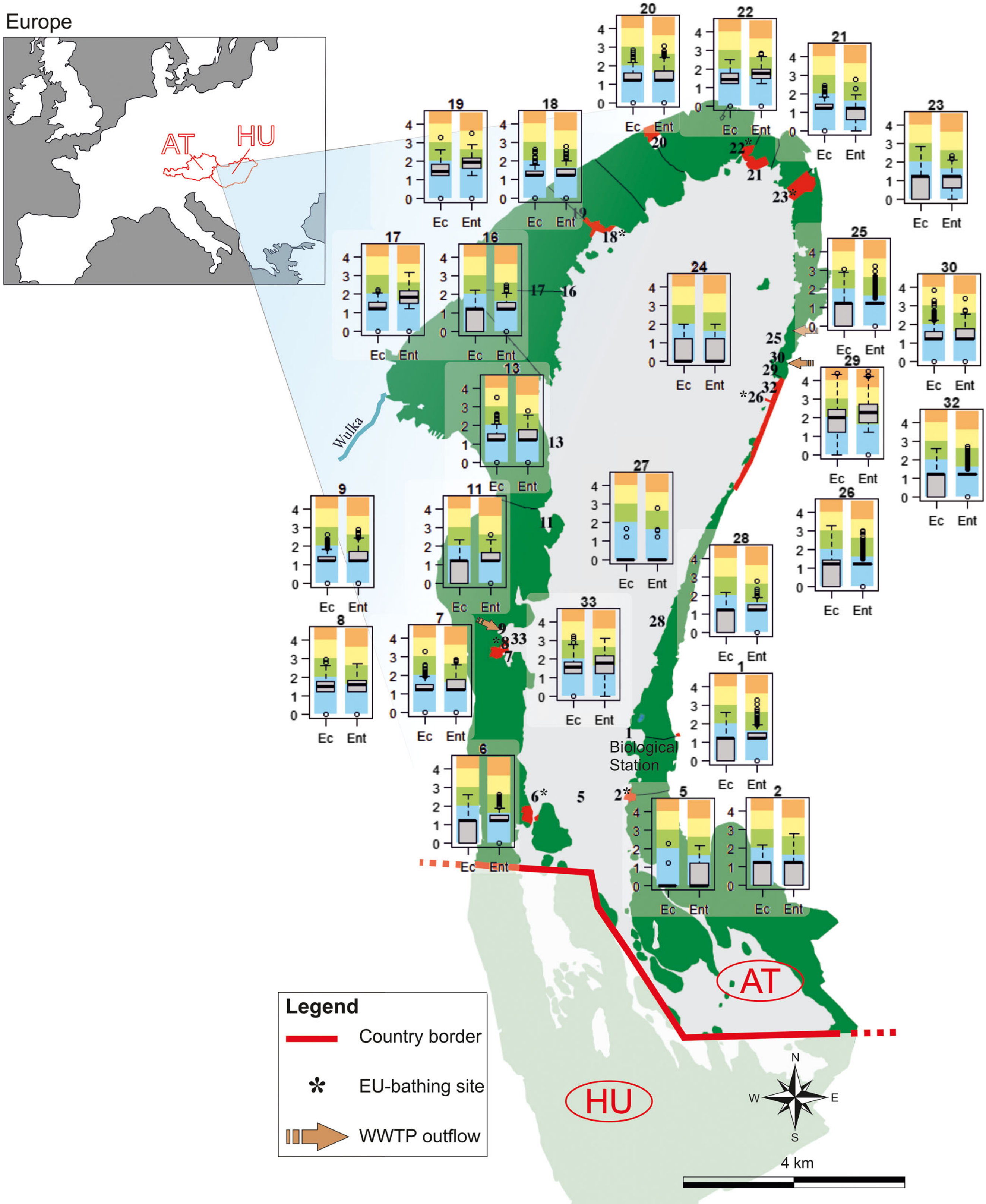


Fig. 1 Boxplots of EC and ENT concentrations at the 26 monitoring sites of Neusiedler See (1992–2013) on a logarithmic scale. Classes of fecal pollution are shown as blue (little), green (moderate), yellow (critical), and orange backgrounds (strong pollution), after Kirschner et al. (2009, 2017). Data falling outside a value of 1.5 times the interquartile range are indicated with a circle (outliers and extreme values); the black lines in the boxes represent the median

discharge data (Q; m s^{-1}) of the Wulka, obtained from the Hydrographical Services of the Province of Burgenland. The whole SFIB and water quality dataset consists of ~ 8700 observations, including ~ 14,700 SFIB and 93,600 water quality measurements (108,300 measurements in total). It should be noted that instead of annual averages, the median or even percentiles could have been used in the analyses. However, in the case of percentiles, the outlying and extreme values would have biased the results, providing a not representative picture of the actual variability of water quality in the lake. In the case of using the medians, a similar pattern would have been obtained as with the used averages since they were almost identical at most of the sites for most of the parameters (e.g. Fig. S1).

Meteorological data

The time series of the meteorological variables (air temperature (°C), precipitation (mm day^{-1}; 7:00 – 7:00 next day), wind speed (m s^{-1} daily avg.), wind direction, hours of sunshine, global radiation (J cm^{-2})) were retrieved from six meteorological stations (Table S2) for the years 1992–2013. The original data were modified only in the case of wind direction to facilitate further data analysis; for details, please see ESM Section "Wind direction conversion." The data sets were obtained from the Austrian Centre for Meteorology and Geodynamics (Zentralanstalt für Meteorologie und Geodynamik) and the Biological Research Institute of Burgenland. The whole meteorological dataset consists of ~ 56,000 measurements.

Methodology and statistical approach

The complexity of the research question, the identification of the driving factors of SFIB variance in a shallow lake, and the determination of the spatiotemporal heterogeneity of the dataset called for a novel methodological approach. For this purpose, the ~ 164,000 individual data points (section "Acquired dataset and preprocessing") were organized into matrices. In this way, principal component analysis could be employed on the SFIB records of all the monitoring sites to identify any hotspots of SFIB variance. Thus, it was also consequently possible to examine the relationship of the obtained principal components—which accounted for most of the SFIB variance in the lake—to the environmental parameters. In addition, via the exploitation of the spatial characteristics of the SFIB data, the spatial representativity of the monitoring network could be assessed through the use of variography (Chilès and Delfiner 2012). The methodology presented serves as an example for other study areas with regard to the question of how similarly large and spatiotemporally heterogeneous datasets should be handled.

Principal component analysis

Principal component analysis (PCA) was used to find the sampling sites with the greatest degree of influence on microbial fecal pollution in the lake over the investigated time period. PCA decomposes the original dependent variables into principal components that explain the original total variance of the dataset in a component-wise monotonically decreasing order. The correlation coefficients between the original parameters and the principal components (PCs) are the factor loadings. They explain the weights of the PCs in the original parameters (Tabachnick and Fidell 2014).

In the present case, the input variables for PCA were the annual averages (1992–2013) for *E. coli* (EC) and intestinal enterococci (ENT) at each sampling site to make the available data comparable due to the highly variable frequency of sampling and the gaps in it. Thus, the obtained factor loadings of the PCs provide information on the importance of each site regarding fecal pollution in the lake over the whole investigated time interval. The sites with factor loadings above 0.7 delineate those areas responsible for most of the microbial fecal pollution variance, and these will be designated "hotspots". The PCs were taken into account based on their scree plots (Cattell 1966) and their eigenvalues, which had to be above 1 (Kaiser 1960).

Subsequently, it was explored whether a relationship exists between the various water quality, hydrographical and meteorological parameters taken as independent variables, and the PCs of EC and ENT as dependent ones. For example, the time series (scores) of the first PCs for EC and ENT were correlated separately with the annual averages of the independent variables measured at those sampling sites which were hotspots of a particular SFIB (Fig. 2). If a significant linear relationship is found to exist, a particular independent variable may then be considered to be a driving factor of microbial fecal pollution over the period 1992–2013; in such a case, the greatest impact would be at the hotspots.

Variogram analysis

Spatial sampling frequency analysis was conducted on the SFIB data for the whole time period investigated with the aim of (1) objectively determining whether the current sampling grid is sufficient or not, and if the grid was insufficient, (2) offering suggestions concerning the recalibration of the network. For this purpose, a geostatistical variogram analysis (Chilès and Delfiner 2012) was carried out, following Hatvani et al. (2017) and Kovács et al. (2012a). The semivariogram describes the spatiotemporal correlation structure of the data. When the distance h between the samples grows (in time or space), the semivariogram values are expected to increase. This implies that the expected difference between the measured responses of the point pairs will also increase (Eq. 1 and Fig. S6).

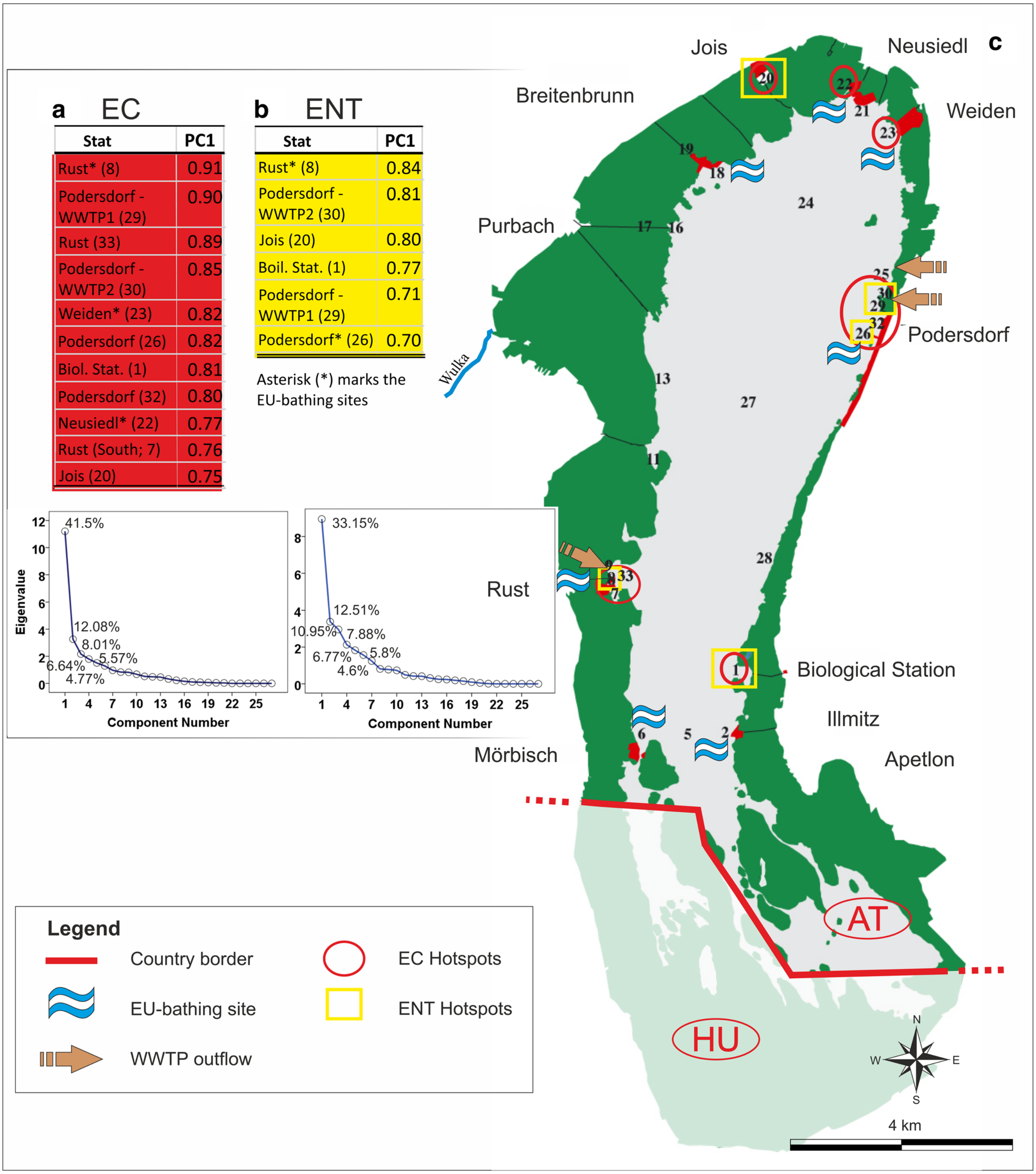


Fig. 2 The loadings and the scree plot of the first PC at the different sampling sites for EC (**a**) and ENT (**b**), and the overall pattern of fecal pollution in the Neusiedler See with the EC (red circle) and ENT (yellow square) hotspots, along with all the sampling sites (**c**). Loadings above the chosen 0.7 threshold are marked in red and yellow for EC (**a**) and ENT (**b**), respectively. The numbers on the map (**a**) and in the brackets (**b**, **c**) represent the sampling sites. The scree plots can be found below tables (**a**) and (**b**) for EC and ENT respectively, with the explained variance of the PCs having an eigenvalue > 1

The core of the semivariogram's delineation is as follows:

Let $Z(x)$ and $Z(x + h)$ be the values of a parameter sampled at points at a distance $|h|$ from each other. The value of h may be measured in time or space for temporal or spatial processes, as required. The semivariogram may then be calculated using the Matheron algorithm (Matheron 1965) (Eq. 1):

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i + h) - Z(x_i)]^2 \quad (1)$$

where $Z(x)$ and $Z(x + h)$ are the values of a parameter sampled at a distance $|h|$ from each other, and $N(h)$ is the number of pairs within the lag interval h.

The most important properties of the semivariogram are (1) the "nugget" C_0, (2) the sill, which is the level at which the semivariogram stabilizes (C = reduced sill + C_0), and (3) the range (a), the distance within which the samples still have an influence on each other as depicted in Fig. S6 (Webster and Oliver 2008). If the semivariogram does not have a rising section, the empirical semivariogram's points will align above the abscissa and parallel to it. A semivariogram like this is called a nugget-effect type of semivariogram, one for which no range can be estimated, i.e., the sampling frequency is insufficient (Hatvani et al. 2017). Samples outside the semivariogram range (be it temporal or spatial) are considered as uncorrelated (Chilès and Delfiner 2012). Thus, the interdependence structure remains undetectable. Therefore, in order to be able to describe the processes, samples should be taken inside the spatial, or—as in this case—temporal range.

In terms of the practical technique, the empirical semivariograms were first derived, then the best-fit theoretical ones modeled in GS+ geostatistical software, as forcefully suggested by Oliver and Webster (2014) and as put into practice by Hatvani et al. (2017). These best-fit model semivariograms were consequently used to provide the weights in the interpolation of the contour maps using Kriging (Chilès and Delfiner 2012; Cressie 1990).

Software used

All statistical computations were performed using the authors' own scripts written in `R 3.2.3` (R Core Team 2018), and Visual Basic for MS Excel 2016, while the geostatistical computations and mapping were carried out with Golden Software Surfer 15 and GS+ 10. For the visualizations of the results, CorelDRAW Graphics Suite X8 and MS Office 2016 were used.

Results

Overall pattern of the variance of the fecal indicator bacteria in the lake

An empirical inspection of the SFIB concentrations at the different sampling sites around the lake revealed substantial differences between the sites (Fig. 1). Open water sites (e.g., sites 5, 24, 27) showed little fecal pollution while occasionally critical to strong pollution levels were observed at specific stations next to settlements (e.g., sites 7, 33, 29, 30). At the open water sites, the averages of EC and ENT values were at their lowest, while the coefficients of variation (CV) were highest. The mean and the median values indicate that this is only due to occasional pollution events, a fact also reflected in the high statistical ranges (Table 1). At the highly polluted sites (e.g., site 29), low CV and high average values were observed, indicating a quasi-continuous SFIB load, while at site 11 both the average and the CV of the SFIB values are low, indicating continuous little to moderate fecal pollution (Table 1).

The 75th quantiles were far from the maximum values at sites near settlements, especially at site 29, where the maximum is nearly twice as high as the 75th quantile, indicating occasional extreme fecal inputs. The average pollution levels at the monitoring locations around site 29 (sites 26, 32) are much lower than at site 29, implying that the fecal pollution from the WWTP effluent is mainly restricted to a small area. At site 26, a relatively high statistical range was observed, most probably due to the close vicinity (< 1 km) of the WWTP effluent (site 29).

In general, the seven EU bathing sites had good water quality with little fecal pollution levels, as reflected in low mean and median values (Table 1, Fig. 1). For a comparison of the data with the guideline values of the EU-BWD, the 90th and 95th percentiles were calculated for all data (period March till October) and for the bathing season only (June–August) using individual observations (Table S3). These data corroborated the good water quality at the seven EU bathing sites. At all bathing sites, the guideline values for good bathing water quality were met, and five of these sites (sites 2, 6, 18, 23, and 26) even complied with the guideline values for excellent water quality. No difference was observed between the two time periods considered (Table S3). Out of the other sampling sites, only four (sites 17, 19, 29, and 33) exceeded the limit values of good bathing water quality, and most of the sites (13 out of 19) complied with the guideline values (EC 2006) for excellent water quality (Table S3).

Main driving factors of fecal pollution in the Neusiedler See

Fecal pollution hotspots in the lake determined using principal component analysis

The first PCs of EC and ENT explained ~42% (Fig. 2a) and ~33% (Fig. 2b) of the SFIB variance in the lake (Fig. 2c), respectively. In all further interpretations, only the first PC was considered in this study, as indicated clearly by the scree plots (Fig. 2a,b). With PCA, the hotspots (PC loadings

Table 1. Descriptive statistics of annual averages of EC and ENT values in Neusiedler See ($\log_{10}$-transformed data). SD stands for standard deviation and CV for coefficient of variation. At all sites and in all cases the minimum value was zero, thus the range determines the maximum values as well. The darker red shades indicate higher, while the darker blue shades lower values. Asterisks (*) mark the EU bathing sites; for more information on the sampling sites, see Tables S1 and S3 and Fig. S1

		EC							ENT						
Name	**Site No.**	**n**	**Mean**	**Median**	**Range**	**q25**	**q75**	**CV**	**n**	**Mean**	**Median**	**Range**	**q25**	**q75**	**CV**
Biological Stat	**1**	620	0.90	1.20	2.62	0	1.20	0.68	621	1.21	1.20	3.26	1.20	1.49	0.46
Ilmitz*	**2**	124	0.67	1.20	2.16	0	1.20	0.96	117	0.86	1.20	2.78	0	1.20	0.77
Open water (South)	**5**	150	0.22	0	2.28	0	0	2.23	150	0.44	0	2.18	0	1.20	1.40
Mörbisch*	**6**	368	0.95	1.20	2.62	0	1.20	0.62	369	1.25	1.20	2.60	1.20	1.49	0.36
Rust (South)	**7**	559	1.27	1.20	3.25	1.20	1.49	0.37	560	1.49	1.20	2.81	1.20	1.76	0.27
Rust*	**8**	553	1.49	1.49	2.93	1.20	1.79	0.30	556	1.60	1.60	2.73	1.20	1.85	0.24
Rust	**9**	543	1.16	1.20	2.62	1.20	1.47	0.47	543	1.39	1.20	2.90	1.20	1.71	0.37
Oggau	**11**	113	0.86	1.20	2.32	0.00	1.20	0.69	113	1.36	1.20	2.58	1.20	1.65	0.37
Oggau	**13**	124	1.17	1.20	3.49	1.20	1.58	0.61	124	1.42	1.20	2.78	1.20	1.77	0.44
Purbach	**16**	116	0.82	1.20	2.19	0.00	1.20	0.80	116	1.24	1.20	2.49	1.20	1.57	0.49
Purbach	**17**	114	1.23	1.20	2.22	1.20	1.55	0.42	114	1.84	1.85	3.18	1.49	2.17	0.29
Breitenbrunn *	**18**	126	1.11	1.20	2.62	1.20	1.47	0.58	126	1.27	1.20	2.78	1.20	1.55	0.45
Breitenbrunn (harbour)	**19**	123	1.48	1.47	3.25	1.20	1.84	0.39	123	1.86	1.92	3.51	1.59	2.15	0.29
Jois	**20**	122	1.23	1.20	2.82	1.20	1.61	0.54	122	1.34	1.20	3.05	1.20	1.72	0.50
Neusiedl (harbour)	**21**	123	1.11	1.20	2.46	1.20	1.47	0.57	123	1.03	1.20	2.78	0.60	1.20	0.63
Neusiedl*	**22**	123	1.49	1.47	2.50	1.20	1.79	0.34	123	1.71	1.76	2.80	1.49	2.00	0.26
Weiden*	**23**	124	0.91	1.20	2.82	0	1.20	0.76	124	1.01	1.20	2.32	0.90	1.20	0.62
open water (North)	**24**	127	0.54	0	1.97	0	1.20	1.16	127	0.49	0	2.00	0	1.20	1.32
Gols channel	**25**	389	1.01	1.20	3.05	0	1.20	0.74	389	1.13	1.20	3.20	1.20	1.20	0.59
Podersdorf*	**26**	554	1.03	1.20	3.25	0	1.47	0.70	556	1.12	1.20	3.00	1.20	1.20	0.57
Open water (middle)	**27**	134	0.14	0	1.70	0	0	2.84	134	0.30	0	2.78	0	0	1.85
Reed belt East	**28**	124	0.83	1.20	2.17	0	1.20	0.79	124	1.12	1.20	2.78	1.20	1.49	0.55
Podersdorf (WWTP1)	**29**	549	1.96	2.00	4.36	1.20	2.46	0.41	553	2.21	2.26	4.48	1.71	2.70	0.33
Podersdorf (WWTP2)	**30**	536	1.25	1.20	3.82	1.20	1.61	0.54	538	1.41	1.20	3.38	1.20	1.79	0.40
Podersdorf	**32**	538	0.91	1.20	2.62	0	1.20	0.74	539	1.02	1.20	2.70	1.20	1.20	0.60
Rust	**33**	287	1.52	1.57	3.20	1.20	1.84	0.38	286	1.78	1.79	3.11	1.20	2.15	0.29

> 0.7) were determined in the lake for both EC and ENT (Fig. 2a, b). It thus became clear that these hotspots account for most of the EC and ENT variability in the lake. It should be noted here that more sites are considered as hotspots in relation to EC than to ENT, but all the sites which were ENT hotspots (Fig. 2b) were also EC hotspots (Fig. 2c). All of these hotspots are without exception located near settlements; none can be found in the reed channels or in the open water (Fig. 2c). The hotspot with the most sites involved was found at Podersdorf (site 29), where all four investigated sites were EC and ENT hotspots, with the exception of site 32 for ENT.

Relationship between the EC and ENT hotspots and the independent variables

To investigate which independent variables might be driving the EC and ENT variance at the hotspots (Fig. 2c), the scores of the first PCs of EC and ENT were correlated with the annual averages of the independent variables' time series, in a way similar to that applied in Hatvani et al. (2015) and Kovács et al. (2015), in both cases using dynamic factor analysis. As far as all the independent water quality and the meteorological parameters go, it was found that water temperature, precipitation at Rust, air temperature, and the number of sunny hours showed a significant negative correlation with the first PC of EC. In the case of ENT, water temperature and precipitation at Rust were negative correlates (Table 2). A positive linear relationship between SFIB and the independent variables was only observed between average- (vv) and maximum wind speed (vv_{max}) recorded at Neusiedl and nitrate-N (Table 2). In the case of ENT, the correlation coefficients did not even approximate $r = \pm 0.7$. The water levels and run-off measured at the Wulka, along with global radiation and other parameters not listed in Table 2, were insignificantly related to the PCs.

Significance levels: $\alpha = 0.05$ (**); $\alpha = 0.1$ (*)

Following the observation that the PCs of the SFIB showed a high and significant degree of correlation with the average and maximum wind speeds, it was investigated to what degree wind speeds from different points of the compass (directional wind speeds) correlate with the PCs of EC and ENT taken separately. By correlating the first PCs for EC and ENT with these directional wind speeds (derived as described in ESM Section "Wind direction conversion"), it was found that with the exception of South, the other directions all have a significant linear relationship with both SFIBs (Table 3).

Significance levels: $\alpha = 0.05$ (**); $\alpha = 0.1$ (*)

Plotting the correlation coefficients of the first PCs of the SFIB and the directional wind speeds (Fig. S5), the most determining directional wind speeds were identified. In the case of EC, the North-West was the most dominant (Fig. 3a), while in the case of ENT, both the North-East and North-West were important, and this can be seen to be related to the SFIB variance in the lake (Fig. 3b), suggesting the prevailing importance of Northerly winds (see ESM Section "Wind direction conversion" for details).

Table 2 Significant correlation coefficients between the independent variables' time series and the EC and ENT PC1 scores

	EC	ENT
W_T	−0.43**	−0.46**
NO_3-N	0.65**	0.41*
T_{air}	−0.43**	–
$Prec_{Rust}$	−0.47**	−0.56**
Sunny hours	−0.50**	–
vv	0.69**	0.41*
vv_{max}	0.4*	–

Table 3 Correlation coefficients between the scores of the first PCs for EC and ENT for the hotspots and the directional wind speed time series attributable to the various wind directions

	EC	ENT
N	0.69**	0.39*
S	0.07	0.11
E	0.68**	0.54**
W	0.82**	0.53**

Spatial sampling frequency estimation

With the application of variography, it was possible to see that the current sampling grid of the sites is too underdeveloped to obtain a spatially continuous geostatistical picture of the whole lake. For ENT, a 1.1 km (Fig. 4a), while for EC a 1.2 km spatial range (Fig. 4b) was modeled. For the two parameters, the kriged maps show a quite similar pattern, dominated by so-called bull's eyes. These indicate that the sites are representative of mostly local phenomena, and further suggesting that interpolation between the sites was insufficient (Chilès and Delfiner 2012). Isolines were plotted only at the multisite hotspots (e.g., Rust, Podersdorf), where the density of the sampling sites was sufficient.

Discussion

The regular input of fecal pollution from WWTPs or receiving waters may continuously impair the water quality of lakes, e.g., the Great Lakes (Nevers et al. 2014) or Lake Geneva in Switzerland (Poté et al. 2009). Short-term fecal pollution events driven by weather phenomena (Patz et al. 2008) or accidental point-source input—e.g., sewer overflow, especially in the case of combined sewer systems (McLellan et al. 2007)—may also have a significant impact on recreational water quality in general, or at specific bathing sites. However, the monitoring of such events is highly difficult due to their ad hoc characteristics (Nevers et al. 2014). In the present study, hotspots and the main drivers of continuous and occasional fecal pollution were identified in order to provide support for management practices in the case of a large shallow Central European lake intensively used for recreational purposes throughout the year (Wolfram et al. 2014). Peer-reviewed studies on the monitoring of long-term (decadal) developments and the drivers of fecal pollution levels in lakes are scarce, e.g., Uejio et al. (2012), Weiskerger and Whitman (2018), and Whitman et al. (1999). It is to be suspected that the main reason for this lies in the lack of academic interest and the difficulties of financing scientific monitoring studies over

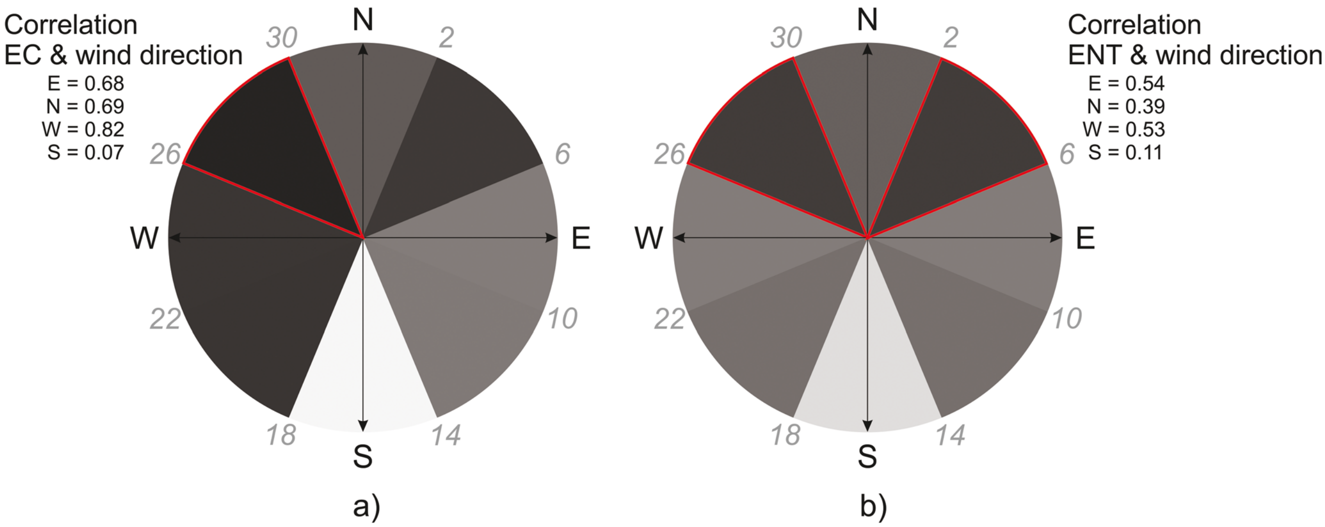


Fig. 3 Overlaid correlations between the annual wind speeds attributed to various wind directions (for details see ESM Section "Wind direction conversion") for EC (**a**) and ENT (**b**). Red triangles mark the wind directions which determine to the greatest degree. The color of the graph is scaled to the magnitude of the correlation coefficient (e.g., r = 0.82 corresponds to 82% black). N, North; E, East, W, West; S, South

and above routine surveillance programs or complex studies on pollution source identification. Such investigations are, however, a prerequisite for effective bathing water management and decision-making, e.g., beach closure (Olyphant 2003; Uejio et al. 2012; Whitman 2003; Whitman et al. 1999). For example, despite the fact that at the largest shallow freshwater lake in Central Europe, Lake Balaton (Istvánovics et al. 2007), there are > 140 official EU bathing sites with monitoring stations (EEA 2018), scientific publications on the assessment of SFIB variance are so far lacking. With regard to general water quality parameters (physical, chemical, and biological), the situation is somewhat better. The monitoring data is assessed to a higher extent, with a significantly larger number of publications available, but still, the amount of information gained from the analyses has not kept pace with the increasing amount of data (Kovács et al. 2012b).

Methodological approaches similar to those employed in the present study can be found; in these, too, the statistical relationship of fecal indicator bacteria and independent meteorological and water quality parameters are assessed to model SFIB variance (Nevers et al. 2014; Nevers and Whitman 2005) and derive a predictive model (Dada and Hamilton 2016; Jones et al. 2013; Uejio et al. 2012). However, these studies do not consider multiple sites together to obtain an overall picture of a lake, and except the long-term data on Geneva Lake, in Wisconsin, USA (Uejio et al. 2012), they cover a much smaller time interval than the present study.

Hotspots of fecal pollution

Descriptive statistics have already revealed that the sampling sites near settlements, especially at Rust and the WWTP outflow at Podersdorf, are the most affected by fecal pollution. By way of contrast, the sites in the open water area of the lake are occasionally affected by fecal pollution, although their background levels are generally low. All the official EU bathing sites at the lake meet the requirements of the EU Bathing Water directive, standards which are less strict than in, e.g., the United States (US EPA 2012) or Canada (Health Canada 2012), due to differences in legislation. According to the US EPA recreational water criteria, the statistical threshold values (which approximate the 90th percentile of the water quality distribution) for *E. coli* and enterococci are 410 and 130 CFU 100 ml^{-1}, respectively (US EPA 2012), in comparison to 900 and 330 CFU 100 ml^{-1} according to the European Bathing water directive (EC 2006).

With the exception of four sites, all other sampling sites (n = 15) complied with the guideline values for good water quality, and 13 sites even displayed excellent water quality (Fig. 1). No difference was observed between the two time periods considered (bathing season June–August versus March–October; Table S3). This indicates that the influence of the bathers during the bathing season is of negligible importance for the water quality of the lake and other influencing factors (see below) prevail.

By assessing the SFIB variability in the lake, hotspots of fecal pollution were identified coinciding with sites at which elevated anthropogenic activity is present (Magyar et al. 2013). The two sub-regions most influential in terms of the variance of EC and ENT in the Austrian part of the lake are located at Rust and Podersdorf (Fig. 2a, b), with multiple sites identified as hotspots. At Rust, only site 9 was not identified as a hotspot, but it should be noted that the loading of site 9 in the first principal component for EC (0.68) is just below the chosen threshold of 0.7. At Podersdorf, all of the sites were identified as EC hotspots except for site 32, located in a bay sheltered from the main currents (Józsa et al. 2008). Site 25, although

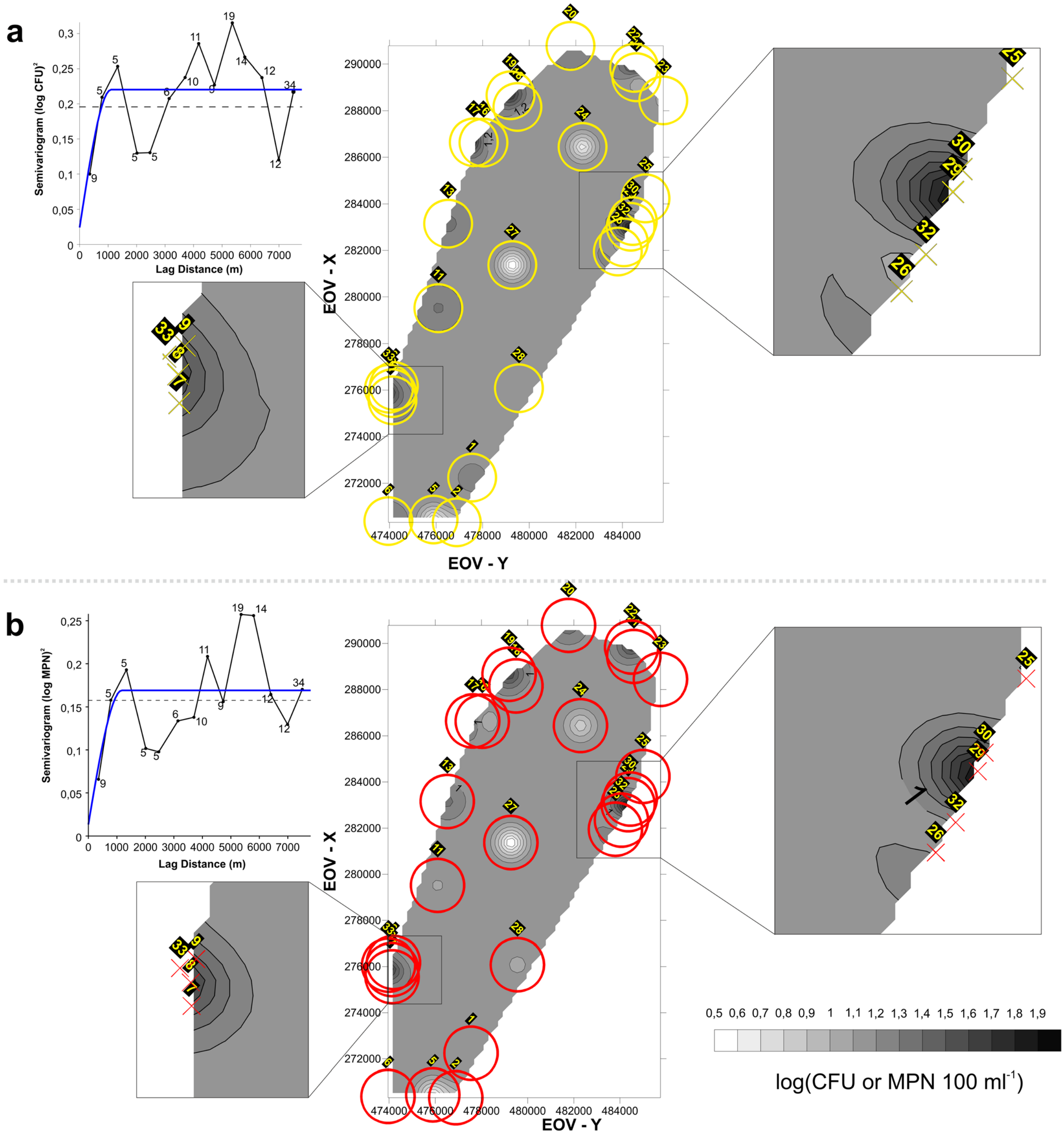


Fig. 4 Empirical (black) and theoretical (red) semivariograms (left) and the kriged isoline maps (right) for **a** ENT and **b** EC. The parameters of the isotropic semivariogram models for ENT were $C_0 = 0.0242$; $C_0 + C = 0.2204$; $a = 1110$ m and for EC: $C_0 = 0.0135$; $C_0 + C = 0.169$; $a = 1180$ m. The semivariograms and the maps represent the 1992–2013 interval and the sites chosen for this study. Colored circles represent the isotropic sampling ranges around the sites for ENT (**a**; yellow) and EC (**b**; red) in accordance with *a* (the range; see "Variogram analysis" section). The horizontal broken line on the semivariogram plots marks the variance, and the numbers indicate the number of data pairs that was used to derive the semivariogram value for a particular bin. For further theoretical information, see Fig. S6. The maps were produced using the Hungarian National Projection System (EOV; NLCO (1971))

located near Podersdorf, was not found to be a hotspot of fecal pollution since it is mainly influenced by the waters of the Golser Channel. Although the channel brings treated wastewater from Mönchhof and Gols, due to the retention ponds on the bank of the Neusiedler See (on the landside edge of the reed belt) and the long residence time through the densely vegetated channel, the impact of the treated wastewater on microbiological water quality is clearly dampened.

The other hotspots (sites 1, 20, 22, and 23; Fig. 2c), despite showing comparable SFIB concentrations to those in the reed belt and/or the open water (e.g., Table 1, Fig. 2c), still have a significant amount of SFIB variance (Fig. 2a,b), in most cases caused by the improper handling of wastewater (Belmont et al. 2004; Burian et al. 2000; Wang et al. 2014). Specifically, in the bay of the Biological Station (site 1) and the channel at Neusiedl (site 22), numerous sailing boats are present during the recreational season. In Jois (site 20) and Weiden (site 23), permanent residences with improper wastewater handling may be the source of the SFIB variance. In the case of Jois, these are located in a rather closed bay, while at Weiden these are found in the corridor of the channel in front of which the measurements are conducted (Fig. 2c).

Meteorological and water quality factors driving microbial fecal pollution in Neusiedler See

Meteorological events are known to have a significant impact on the microbial pollution status of surface waters (Patz et al. 2008). Storm events, sewer overflow, or increased surface run-off due to heavy precipitation events lead to short-term increases in fecal indicator concentrations. Similar phenomena were expected in the case of the Neusiedler See, and all the more so, as the lake is very shallow and dilution of fecal pollution is less effective. In fact, some significant correlations were found between SFIB variance and meteorological parameters. On the one hand, the observed negative correlations between the SFIB variance and particular meteorological parameters (Table 2) imply that if there is no wind/precipitation, the water and air temperature are more likely to be higher, along with a greater chance of an elevated number of sunny hours. In these cases, peaks in SFIB presence are less likely to occur, as was also found to be so in the case of the Great Lakes (Nevers et al. 2014). On the other hand, SFIB correlated with wind speed in a positive way (Tables 2 and 3). For both EC and ENT, the Northwesterly wind direction was found to be the most determining (Fig. 3). This concurs with the wind directions (N, NNW, NW, SSE) observed to be dominant in the joint campaign of Józsa et al. (2008), consisting of on-site measurements. In the case of the shallow waters at the Podersdorf beach, wave action caused by the onshore winds results in the resuspension of the sediment. Recent studies along the coast of Lake Michigan indicate that the bacteria concentration has grown so high in the sand of the beaches that wave action can stir up the sediment and temporarily release bacteria into the water. This, in turn, resulted in high concentrations of *E. coli* (Olyphant 2003; Whitman 2003). Such scenario may also happen along the Podersdorf beach as well, though unfortunately no recent data are available on the bacteria concentrations in the sediments (Dokulil 1975).

In the case of heavy rains, mostly accompanied by NW winds, the WWTP effluent channel at Podersdorf overflows, accounting for SFIB peaks at the bathing sites. Urban stormwater has emerged as a major contributor of bacterial and chemical pollution in watersheds (McLellan and Jensen 2003), especially in the case of combined sewer systems, where heavy rainfall exceeds the holding capacity of the system and a mixture of sanitary sewage and stormwater is discharged into the lake. In Podersdorf, for example, 97% are part of a combined sewer system (Wolfram et al. 2014).

In the case of the second most important hotspot, Rust, it had already been shown previously that in terms of general water quality data, on account of the increased recreational usage and possibly improper handling of domestic sewage water in the area, private weekend huts cause pollution patterns different from, e.g., the open water or the reed belt (Magyar et al. 2013). Although the obtained principal components in the present study represent the whole lake, the SFIB variance at Rust was highly determining (Fig. 2a, b). It is assumed that this is directly related to the weekend huts, as also in the case of the general water quality parameters (Magyar et al. 2013). Thus, if it rains, anthropogenic presence/activity decreases, lowering the amount of fecal input at Rust, manifested in the significant negative relationship between precipitation at Rust and the EC and ENT variance (Table 2).

The significant positive relationship between nitrate concentrations and SFIB variance may be due to diffuse pollution from livestock and horses (Canter 1996) in the area. No significant correlations were found between the discharge of the main tributary, the Wulka, and of the water level, with the fecal pollution at the hotspots. Despite the fact that the Wulka transports the wastewater from several WWTPs along the river, its influence on microbial lake water quality on an annual scale is negligible. This can be explained by the fact that the water of the Wulka passes through a 3-km-wide reed stand before reaching the open water of the lake (Dinka et al. 2016). Also, annual water level fluctuations seem to be unimportant despite the fact that a concentration of fecal pollution at lower water levels may be assumed. In contrast, water level changes may be relevant at the short time scale when strong winds induce increased wave action leading to resuspension of sediments and—in case of westerly winds at the Wulka Delta—mobilization of fecal bacteria retained in the red belt into the lake.

Differences in the correlation patterns (more EC than ENT hotspots were found) can be explained by the greater persistence of enterococci in comparison to *E. coli*, especially in sub-saline environments (Jin et al. 2004; Kirschner et al. 2014). Thus, EC variability (input and die-off) is more determined by meteorological and local effects (~42% explained by the first PC) than ENT (~33%). Consequently, further investigations are needed to determine the main controlling factors of ENT variability at the sites which are solely EC hotspots.

Geostatistical assessment of SFIB variance

Overall, it can be concluded that with an approximately 1 km spatial range, interpolation was insufficient in the case of distant sites with the current sampling grid, and that in general the sites are mostly representative of local phenomena. Using a rigorous method to determine homogeneous groups of sampling sites (Kovács et al. 2014), concurring results were obtained with the present ones. In a particular study based on multiple water quality variables, most of the sampling sites of the lake formed separate groups. However, in the areas of the hotspots determined and discussed in the present paper, homogeneous groups of multiple sampling sites were observed (Kovács et al. 2014). Thus, the vicinity of the main hotspots (e.g., Podersdorf and Rust)—where multiple sites were located within the spatial range—was represented by "evaluable" isolines, not only bull's eyes (Fig. 4). In addition, a similar pattern was observed on the isoline map of the scores of the first principal component of general water quality variables (mainly inorganic ones), clearly indicating the anthropogenic influence on the water quality of the lake at Rust and by the Podersdorf WWTP (Magyar et al. 2013).

The extension of the influence of the fecal pollution hotspots (Fig. 4) is comparable to those modeled in the surface sediments of Bay of Vidy on Lake Geneva (Switzerland), where the WWTP's outlet was close to the shore in a shallow area (Poté et al. 2008). From a methodological aspect, we presume that the maps from Lake Geneva were also modeled using variography and kriging. However, due to the lack of a documented semivariogram model, the actual representativity and the spatial correlation structure of their data (Poté et al. 2008) cannot be evaluated, unlike in the present case. In cases where the subject of the spatial modeling was bacteriophage tracer tests—considered to be consistent with the spatiotemporal variability of SFIB (Goldscheider et al. 2007)—interpolation with kriging was not possible due to the lack of homogeneous mixing processes (Goldscheider et al. 2007; Wanninkhof et al. 2005). At the Neusiedler See hotspots, the number of investigated sampling locations was lower than in the previously mentioned studies. Nevertheless, interpolation was sufficient, probably due to the steady-state flow patterns in the lake ($v = 10$ m s^{-1}) in the area (Józsa et al. 2008). This suggests that from a scientific point of view, a denser and planned campaign—using a quasi-equidistant grid—should be commenced to map the SFIB variance over the whole lake, for example, a couple of times prior to and then after storm events in the course of one bathing season. In this way, the spatial SFIB variability over the whole lake, but especially at the hotspots, could be sufficiently determined. Then, from a practical point of view, the obtained results could be used to plan and set up an economically optimized/recalibrated monitoring system, which would reliably represent SFIB variance over the whole lake, at the same time allowing for a special focus on the problematic areas with alarming SFIB variance.

Conclusions

For lakes intensively used for recreation, specific microbiological water quality targets have to be met. Thus, spatiotemporally representative surveillance needs to be conducted spanning a set of various biotic and abiotic water quality parameters. However, such surveillance may be in itself insufficient to elucidate the main drivers of fecal pollution events in a complex lake/watershed ecosystem. The in-depth analysis of the microbial fecal pollution data at hand coupled with additional environmental variables determined at a large number of sampling sites is needed. In the present study, with the application of multivariate statistical and geostatistical methods, the hotspots of standard fecal indicator bacteria (SFIB) variability were identified and delineated in space in a highly important recreational shallow lake in Europe. The presence of the hotspots was demonstrated to be attributable to the presence of extensive anthropogenic activity, such as the emission of effluents from wastewater treatment plants or the improper handling of communal sewage. At sites located remotely from the fecal pollution hotspots, while continuously low SFIB levels were to be observed, peaks of SFIB presence were noted, underlining the affected state of the whole lake. Thus, a greater effort is required to reduce fecal pollution inputs at the hotspot sites.

Based on the geostatistical analysis, it is suggested that an intense campaign be initiated, both in time and space, in order to be able to map the different areas/habitats of the lake more representatively in terms of SFIB. Were any such campaign to be planned, the prevailing wind directions (mostly northerly) should be taken into account. With the use of variography, the present grid explicitly indicates the areas as yet uncovered by the current network of sampling sites. On the basis of this information, a sustainable, economically optimized monitoring program could be set up, surpassing the current surveillance activities at the seven Austrian EU bathing sites with the aim of improving recreational water quality over the whole lake.

In addition to information on the extent and spatiotemporal patterns of recent fecal pollution based on *E. coli* and intestinal enterococci concentrations, more persistent fecal pollution indicators and microbial source tracking techniques—e.g., Edge et al. (2010), Kirschner et al. (2017), and Reischer et al. (2011)—should be applied to elucidate the patterns and identify the main drivers of and contributors to fecal pollution in complex lake/watershed ecosystems. In the area under consideration, the implementation of *Clostridium perfringens* or coliphages into the monitoring concept and the application of human-, horse-, pig-, bird-, and ruminant-associated microbial fecal pollution markers for source tracking purposes would significantly improve the understanding of the observed complex pollution patterns and efficiently support bathing-water management practices.

Acknowledgements Open access funding provided by Medical University of Vienna. The authors thank Paul Thatcher for thoroughly proofreading and language editing a former version of this manuscript. We also thank Prof. József Kovács from Eötvös Loránd University for the stimulating discussions. The study was financed by the Austrian Provincial Government Burgenland (Amt der Burgenländischen Landesregierung, Abteilung 5). Additional support was provided by the MTA "Lendület" program (LP2012-27/2012) and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, the Hungarian Ministry of Human Capacities (NTP-NFTÖ-17), and the Austro Hungarian Action Fund (97öu1).

References

Bartram J, Richard B, World Health Organization, United Nations Environment Programme (1996) Water quality monitoring: a practical guide to the design and implementation of freshwater quality studies and monitoring programs. E & FN Spon, London

Bedri Z, Corkery A, O'Sullivan JJ, Deering LA, Demeter K, Meijer WG, O'Hare G, Masterson B (2016) Evaluating a microbial water quality prediction model for beach management under the revised EU bathing water directive. J Environ Manag 167:49–58

Belmont MA, Cantellano E, Thompson S, Williamson M, Sánchez A, Metcalfe CD (2004) Treatment of domestic wastewater in a pilot-scale natural treatment system in Central Mexico. Ecol Eng 23:299–311

Burian SJ, Nix SJ, Pitt RE, Durrans SR (2000) Urban wastewater Management in the United States: past, present, and future. J Urban Technol 7:33–62

Canter LW (1996): Nitrates in groundwater. CRC press

Cattell RB (1966) The scree test for the number of factors. Multivar Behav Res 1(2):245–276

Chapman DV, Bradley C, Gettel GM, Hatvani IG, Hein T, Kovács J, Liska I, Oliver DM, Tanos P, Trásy B, Várbíró G (2016) Developments in water quality monitoring and management in large river catchments using the Danube River as an example. Environ Sci Pol 64:141–154

Chilès J-P, Delfiner P (2012) Geostatistics. Wiley, Canada

Convention on Wetlands 1971: Convention on Wetlands of International Importance especially as Waterfowl Habitat, Ramsar, Iran, 2 February 1971. UN Treaty Series no. 14583. As amended by the Paris Protocol, 3 December 1982, and Regina Amendments, 28 May 1987, UN, Brussels, Ramsar, Iran

Cressie N (1990) The origins of kriging. Math Geol 22:239–252

Dada AC, Hamilton DP (2016) Predictive models for determination of E. coli concentrations at inland recreational beaches. Water Air Soil Pollut 227:347

Dinka M, Kiss A, Magyar N, Ágoston-Szabó E (2016): Effects of the introduction of pre-treated wastewater in a shallow lake reed stand. Open Geosciences, pp 62

Dokulil MT (1975) Bacteria in the water and mud of Neusiedlersee (Austria). Symp Biol Hung 15:135–140

Dokulil MT (2014) Environmental impacts of tourism on lakes. In: Ansari AA, Gill SS (eds) Eutrophication: causes, consequences and control: volume 2. Springer Netherlands, Dordrecht, pp 81–88

Dokulil MT (2017) Peri-Alpine Lakes in the Anthropocene: deterioration and rehabilitation—an Austrian success story. Acta ZooBot Austria 154:1–53

EC (2000) Directive 2000/60/European Commission of the European Parliament and of the Council of 23 October 2000 establishing a framework for community action in the field of water policy. Off J Eur Communities L 327:1–72

EC (2006) Directive 2006/7/EC of the European Parliament and of the Council of 15 February 2006 concerning the management of bathing water quality and repealing Directive 76/160/EEC. Off J Eur Communities L 64:1–37

Edge TA, Hill S, Seto P, Marsalek J (2010) Library-dependent and library-independent microbial source tracking to identify spatial variation in faecal contamination sources along a Lake Ontario beach (Ontario, Canada). Water Sci Technol 62:719–727

EEA (2018): State of bathing waters. European Environment Agency, Brussels

Goldscheider N, Haller L, Poté J, Wildi W, Zopfi J (2007) Characterizing water circulation and contaminant transport in Lake Geneva using bacteriophage tracer experiments and limnological methods. Environ Sci Technol 41:5252–5258

Hatvani IG, Clement A, Kovács J, Kovács IS, Korponai J (2014a) Assessing water-quality data: the relationship between the water quality amelioration of Lake Balaton and the construction of its mitigation wetland. J Great Lakes Res 40:115–125

Hatvani IG, Magyar N, Zessner M, Kovács J, Blaschke AP (2014b) The water framework directive: can more information be extracted from groundwater data? A case study of Seewinkel, Burgenland, eastern Austria. Hydrogeol J 22:779–794

Hatvani IG, Kovács J, Márkus L, Clement A, Hoffmann R, Korponai J (2015) Assessing the relationship of background factors governing the water quality of an agricultural watershed with changes in catchment property (W-Hungary). J Hydrol 521:460–469

Hatvani IG, Leuenberger M, Kohán B, Kern Z (2017) Geostatistical analysis and isoscape of ice core derived water stable isotope records in an Antarctic macro region. Polar Science 13:23–32

Health Canada, Water AaCCB, Healthy Environments and Consumer Safety Branch (2012): Guidelines for Canadian recreational water quality. Health Canada, Ottawa, Ontario, Canada, 161 pp

Istvánovics V, Clement A, Somlyódy L, Specziár A, G- Tóth L, Padisák J (2007) Updating water quality targets for shallow Lake Balaton (Hungary), recovering from eutrophication. In: Qin B, Liu Z, Havens K (eds) Eutrophication of shallow lakes with special reference to Lake Taihu, China. Springer Netherlands, Dordrecht, pp 305–318

Jin G, Englande A, Bradford H, H-w J (2004) Comparison of E. coli, enterococci, and fecal coliform as indicators for brackish water quality assessment. Water Environ Res 76:245–255

Jones RM, Liu L, Dorevitch S (2013) Hydrometeorological variables predict fecal indicator bacteria densities in freshwater: data-driven methods for variable selection. Environ Monit Assess 185:2355–2366

Józsa J, Krámer T, K. H, Napoli E, Sarkkula J 2008: Wind-induced hydrodynamics and sediment transport of Lake Neusiedl—Hungarian–Austrian–Finnish research cooperation from lake-wide to bay-wide scale. Budapest University of Technology and Economics, Department of Hydraulic and Water Resources Engineering, Budapest

Kaiser HF (1960) The application of electronic computers to factor analysis. Educ Psychol Meas 20(1):141–151

Kirschner AKT, Kavka GG, Velimirov B, Mach RL, Sommer R, Farnleitner AH (2009) Microbiological water quality along the Danube River: integrating data from two whole-river surveys and a transnational monitoring network. Water Res 43:3673–3684

Kirschner A, Herzig A, Palffy K, Wolfram G (2014): Mikrobiologie und Hygiene. In: Wolfram G, Déri L , Zech S (Editors), Strategic study

for lake Neusiedler See—phase 1. Austrian–Hungarian Water Commission, Vienna—Szombathely, pp. 246
Kirschner AKT, Reischer GH, Jakwerth S, Savio D, Ixenmaier S, Toth E, Sommer R, Mach RL, Linke R, Eiler A, Kolarevic S, Farnleitner AH (2017) Multiparametric monitoring of microbial faecal pollution reveals the dominance of human contamination along the whole Danube River. Water Res 124:543–555
Kovács J, Korponai J, Székely Kovács I, Hatvani IG (2012a) Introducing sampling frequency estimation using variograms in water research with the example of nutrient loads in the Kis–Balaton water protection system (W Hungary). Ecol Eng 42:237–243
Kovács J, Tanos P, Korponai J, Székely IK, Gondár K, Gondár-Sőregi K, Hatvani IG (2012b): Analysis of water quality data for scientists. In: Kostas Voudouris , Voutsa D (Editors), Water quality monitoring and assessment. InTech, pp. 65–94
Kovács J, Kovács S, Magyar N, Tanos P, Hatvani IG, Anda A (2014) Classification into homogeneous groups using combined cluster and discriminant analysis. Environ Model Softw 57:52–59
Kovács J, Márkus L, Szalai J, Kovács IS (2015) Detection and evaluation of changes induced by the diversion of river Danube in the territorial appearance of latent effects governing shallow-groundwater fluctuations. J Hydrol 520:314–325
Kubu G, Krámer T, Szilágyi J (2014): Hydrology and water management. In: Wolfram G, Déri L , Zech S (Editors), Strategic study for lake Neusiedler See—phase 1. Austrian–Hungarian Water Commission, Vienna—Szombathely, pp. 246
Magyar N, Hatvani IG, Székely IK, Herzig A, Dinka M, Kovács J (2013) Application of multivariate statistical methods in determining spatial changes in water quality in the Austrian part of Neusiedler See. Ecol Eng 55:82–92
Matheron G (1965): Les Variables régionalisées et leur estimation: une application de la théorie des fonctions aléatoires aux sciences de la nature. Masson et Cie Luisant-Chartres, impr. Durand, Paris
McLellan S, Jensen E (2003) Lake Michigan beaches: urban stormwater and water quality advisories. LakeLine 23:27–29
McLellan SL, Hollis EJ, Depas MM, Van Dyke M, Harris J, Scopel CO (2007) Distribution and fate of Escherichia coli in Lake Michigan following contamination with urban stormwater and combined sewer overflows. J Great Lakes Res 33:566–580
MEA (2005) Ecosystems and human well-being: synthesis. Island Press, Washington
Nevers MB, Whitman RL (2005) Nowcast modeling of Escherichia coli concentrations at multiple urban beaches of southern Lake Michigan. Water Res 39:5250–5260
Nevers MB, Byappanahalli MN, Edge TA, Whitman RL (2014) Beach science in the Great Lakes. J Great Lakes Res 40:1–14
NLCO (1971) Projection guideline for the application of the uniform National Projection System. In: Budapest
Oliver MA, Webster R (2014) A tutorial guide to geostatistics: computing and modelling variograms and kriging. CATENA 113:56–69
Olyphant G (2003) Beach closure decisions: the case for developing and using prediction models. LakeLine 23:19–22
Patz JA, Vavrus SJ, Uejio CK, McLellan SL (2008) Climate change and waterborne disease risk in the Great Lakes region of the U.S. Am J Prev Med 35:451–458
Poté J, Haller L, Loizeau J-L, Garcia Bravo A, Sastre V, Wildi W (2008) Effects of a sewage treatment plant outlet pipe extension on the distribution of contaminants in the sediments of the bay of Vidy, Lake Geneva, Switzerland. Bioresour Technol 99:7122–7131
Poté J, Haller L, Kottelat R, Sastre V, Arpagaus P, Wildi W (2009) Persistence and growth of faecal culturable bacterial indicators in water column and sediments of Vidy Bay, Lake Geneva, Switzerland. J Environ Sci 21:62–69
R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
Reischer GH, Kollanur D, Vierheilig J, Wehrspaun C, Mach RL, Sommer R, Stadler H, Farnleitner AH (2011) Hypothesis-driven approach for the identification of fecal pollution sources in water resources. Environ Sci Technol 45:4038–4045
Sommer R, Blaschke A, Farnleitner A, Kirschner A 2018: Nachhaltiges Gesamtkonzept zum Management des mikrobiologisch-hygienischen Zustandes des Neusiedler Sees unter Berücksichtigung der Siedlungswasser-wirtschaft der Umlandgemeinden. Phase 1: "Pilotstudie". Report to the Austrian Federal Ministry of Federal Ministry of Agriculture, Environment, Forestry and Water Management, Vienna
Strobl RO, Robillard PD (2008) Network design for water quality monitoring of surface freshwaters: a review. J Environ Manag 87:639–648
Tabachnick BG, Fidell LS (2014): Using multivariate statistics. Pearson Education Limited, Essex, England
Uejio CK, Peters TW, Patz JA (2012) Inland lake indicator bacteria: long-term impervious surface and weather influences and a predictive Bayesian model. Lake Reservoir Manag 28:232–244
US_EPA UEPA 2012: Recreational Water Quality Criteria
Wang H, Wang T, Zhang B, Li F, Toure B, Omosa IB, Chiramba T, Abdel-Monem M, Pradhan M (2014) Water and wastewater treatment in Africa—current practices and challenges. CLEAN—Soil, Air, Water 42:1029–1035
Wanninkhof R, Sullivan KF, Dammann WP, Proni JR, Bloetscher F, Soloviev AV, Carsey TP (2005) Farfield tracing of a point source discharge plume in the Coastal Ocean using sulfur hexafluoride. Environ Sci Technol 39:8883–8890
Ward RC, Loftis JC, McBride GB (1986) The "data-rich but information-poor" syndrome in water quality monitoring. Environ Manag 10: 291–297
Webster R, Oliver MA (2008): Geostatistics for environmental scientists. Geostatistics for environmental scientists. John Wiley & Sons, Ltd, 330 pp
Weiskerger CJ, Whitman RL (2018) Monitoring E. coli in a changing beachscape. Sci Total Environ 619-620:1236–1246
Whitman RL (2003) Monitoring lakes for recreational water quality. LakeLine 23:16–18
Whitman RL, Nevers MB, Gerovac PJ (1999) Interaction of ambient conditions and fecal coliform bacteria in southern Lake Michigan beach waters: monitoring program implications. Nat Areas J 19: 166–171
WHO (2003) Guidelines for safe recreational water environments. Volume 1: coastal and fresh waters. WHO, Geneva
Wolfram G, Déri L, Zech S (2014) Strategic study for lake Neusiedler See—phase 1. Austrian–Hungarian Water Commission, Vienna—Szombathely, 246 pp
Wymer LJ (2007) Statistical framework for recreational water quality criteria and monitoring, 65. John Wiley & Sons

istvan60_104_23

Polar Science 13 (2017) 23–32

Contents lists available at ScienceDirect

Polar Science

journal homepage: https://www.evise.com/profile/#/JRNL_POLAR/login

Geostatistical analysis and isoscape of ice core derived water stable isotope records in an Antarctic macro region

István Gábor Hatvani [a,*], Markus Leuenberger [b,c], Balázs Kohán [d], Zoltán Kern [a]

[a] Institute for Geological and Geochemical Research, Research Centre for Astronomy and Earth Sciences, MTA, Budaörsi út 45., H-1112 Budapest, Hungary
[b] Division of Climate and Environmental Physics, Physics Institute, University of Bern, Sidlerstrasse 5, CH-3012, Bern, Switzerland
[c] Oeschger Centre for Climate Change Research, Falkenplatz 16, CH-3012, Bern, Switzerland
[d] Department of Environmental and Landscape Geography, Eötvös Loránd University, Budapest, Hungary

ARTICLE INFO

Article history:
Received 14 November 2016
Received in revised form
6 April 2017
Accepted 7 April 2017
Available online 11 April 2017

Keywords:
$\delta^{18}O$ & $\delta^{2}H$ records
Isoscape
Polar precipitation
Variogram analysis

ABSTRACT

Water stable isotopes preserved in ice cores provide essential information about polar precipitation. In the present study, multivariate regression and variogram analyses were conducted on 22 δ^2H and 53 $\delta^{18}O$ records from 60 ice cores covering the second half of the 20th century. Taking the multicollinearity of the explanatory variables into account, as also the model's adjusted R^2 and its mean absolute error, longitude, elevation and distance from the coast were found to be the main independent geographical driving factors governing the spatial $\delta^{18}O$ variability of firn/ice in the chosen Antarctic macro region. After diminishing the effects of these factors, using variography, the weights for interpolation with kriging were obtained and the spatial autocorrelation structure of the dataset was revealed. This indicates an average area of influence with a radius of 350 km. This allows the determination of the areas which are as yet not covered by the spatial variability of the existing network of ice cores. Finally, the regional isoscape was obtained for the study area, and this may be considered the first step towards a geostatistically improved isoscape for Antarctica.

1. Introduction

Due to the increasing interest in the understanding of past global changes, additional and complementary information about past climates is needed. Ice cores play an important role in relation to this issue (EPICA, 2006; NGRIP, 2004; Wolff et al., 2010). For instance, the water stable isotope characteristics stored in them hold crucial information concerning the precipitation they were formed from. The isotopic composition of precipitation, in turn, gives insights into (i) the origin of the water vapor, (ii) the conditions during condensation, and (iii) those during precipitation (Araguás-Araguás et al., 2000; Dansgaard, 1964; Merlivat and Jouzel, 1979). Ice cores can yield information about past climates ranging in time-scale from the seasonal (Hammer, 1989; Kuramoto et al., 2011) up to several hundred millennia (EPICA, 2004), and provide relevant indications about the large-scale dynamics of the Earth's climatic system (Jouzel, 2013). By integrating the knowledge gained from studying stable isotopes in ice cores into global circulation models, a more detailed picture can be obtained of the climatic factors driving temporal water isotope variability (Werner and Heimann, 2002).

However, dealing with stable isotope data from ice cores in Antarctica is a challenging task, since the spatial availability of cores is sparse and highly variable over the continent (IPICS, 2006; Masson-Delmotte et al., 2008; Steig et al., 2005). Apart from process-based modeling, interpolation is therefore one of the only means available to make estimations between locations for which data are available (Rotschky et al., 2007; Wang et al., 2010).

Interpolated maps representing the global distribution of water stable isotopes in precipitation have been developed (Terzer et al., 2013; van der Veer et al., 2009). These, however, do not cover Antarctica. The only product that maps the spatial distribution of stable isotopic composition in Antarctic surface snow (Wang et al., 2010) neglects the shelf areas. Of these regions, the Filchner-Ronne-, Riiser-Larsen and Fimbul ice shelves cover a fair portion of the area investigated in the present study.

* Corresponding author. Institute for Geological and Geochemical Research, Research Centre for Astronomy and Earth Sciences, Hungarian Academy of Sciences, H-1112 Budapest, Budaörsi út 45 Hungary.

E-mail addresses: hatvaniig@gmail.com (I.G. Hatvani), leuenberger@climate.unibe.ch (M. Leuenberger), balazs.kohan@gmail.com (B. Kohán), zoltan.kern@gmail.com (Z. Kern).

URL: http://paleoclimate2ka.hu/?lang=en

http://dx.doi.org/10.1016/j.polar.2017.04.001
1873-9652/

Isoscapes are predictive models that estimate the local isotopic composition of environmental materials as a function of observed local and/or extralocal environmental variables (Bowen, 2010). The horizontal and vertical resolution of isotope enabled global circulation models (GCMs) are steadily improving (e.g. Werner and Heimann, 2002; Xi, 2014), such that isotope enabled GCMs using resolutions previously only attainable in regional models are now available (Sjolte et al., 2011; Werner et al., 2011). In the settings where station based precipitation stable isotope records are available, these are naturally the primary inputs to evaluate the performance of isotope enabled circulation models (Lachniet et al., 2016; Sturm et al., 2005). However, gridded products of precipitation stable isotopes (e.g. isoscapes) can be used as additional benchmarks when observations are missing to assess the global/regional circulation models' effectiveness in replicating observed/interpolated data representing the hydrological cycle and its isotopic counterparts.

The aims of this study were (i) to determine the geographic factors driving the stable isotope variability in a chosen Antarctic macro region; (ii) to assess the spatial continuity properties (variograms) of the stable isotope records, an absolute necessity for geostatistical mapping (Herzfeld, 2004), and (iii) to determine the regional isoscape for ice core derived stable isotope records.

Variogram analysis was used in the hope that it would reveal those areas insufficiently represented by the current set of ice cores, giving an indication of where their spatial coverage might be increased and reveal the spatial dependence structure of the stable isotope records. In addition, variography is vital for kriging (Cressie, 1990; Oliver and Webster, 2014; van der Veer et al., 2009), an "optimal" interpolation which is then employed in the study to estimate the covariances to the highest degree of accuracy possible before mapping. Consequently, the derived isoscape (Bowen, 2010) will be able to describe the spatial distribution of isotopes in the region in a representative way.

Worthy of mention is the fact that the aims of this study are in close agreement with the goals of the International Partnerships in Ice Coring Sciences initiative, since the regional nature of climate and climate forcing requires data from a geographically extensive area. In addition, in order to be able to interpret the water stable isotope records from the past 2 ky of ice cores precisely, these have to be supplemented by additional shorter cores for validation (IPICS, 2006).

2. Materials and methods

2.1. Description of the study area and the used dataset

The Antarctic study area (Latitude: 71°S, 83°S; Longitude: 61°W, 12°E; Fig. 1) covering about 2.6×10^6 km^2 in the Atlantic sector, was chosen on account of the relatively high abundance of available ice core derived water stable isotope records, and the fact that it disposes of numerous deep ice cores, which have played and continue to play an important role in paleoclimatology. The region is considered to be diverse from both the topographic and glacioclimatologic perspectives, as well (Graf et al., 1994; Oerter et al., 2000; Rotschky et al., 2007), with areas of low elevation (e.g. the Coastal Dronning Maud Land, Ronne Ice Shelf etc.) at sea level, and significantly higher ones (e.g. the Central Dronning Maud Land > ~2500 m a.s.l.). Field observations have shed light on an atypical continental precipitation distribution obtaining in the region, in which the accumulation and mean air temperature decrease with distance from the shoreline and with the increase in elevation (Vaughan et al., 1999). The difference in accumulation between the highly elevated inland regions and the coast may be as great as a factor of six (Graf et al., 1994; Oerter et al., 2000), and vary by up to e.g. 500 kg m^{-2} a^{-1} over a distance of <3 km in certain areas of the Western Dronning Maud Land (Rotschky et al., 2007); for details see Table S1.

2.2. Dataset used

The data used were acquired from open access data repositories (NOAA (2014); PANGAEA (2014)) and the corresponding research groups (Divine et al., 2009; Naik et al., 2010). Altogether, an array of 22 δ^2H (Figs. S1a and b) and 53 δ^{18}O (Figs. S1c and d) records was assembled from 60 ice cores spanning various time intervals. In the compiled ice core derived water isotope database, isotope abundances are expressed as per mil (‰), differences from the V-SMOW standard (Coplen, 1994) using the δ notation, $\delta X = [(R_{sample}/R_{standard}) - 1] \times 1000$, where X is ^{2}H or ^{18}O, R_{sample} is the sample ^{2}H/^{1}H or ^{18}O/^{16}O ratio, and $R_{standard}$ is the ^{2}H/^{1}H or ^{18}O/^{16}O ratio of the standard. The longest time interval spanned was almost a millennium (Fig. S1d), while the shortest covered only a couple of years (Fig. S1c).

The study was restricted to the period 1970–1988, corresponding to 44 δ^{18}O, and from 1970 to 1989 with 22 δ^2H records before pre-processing and filtering. In this way, both the time span and the available number of records were maximized. By choosing the higher number of cores against the longer timescale, the possibility of better signal replication arose, as emphasized e.g by Jones et al. (2009). In addition, there were five ice cores (c5, c7, c9, c11 and c13) which only had δ^2H records; these were converted to δ^{18}O using the regional δ^2H - δ^{18}O relation established (Fig. S2) based on five neighboring cores with both δ^2H and δ^{18}O records, (for details see SOM). Note that in one special case, the δ^{18}O and δ^2H records of two cores spaced only 6 km apart, namely, c48 & c49 NM01C82 _04 (B04) and NM02C02_02 (FB0202) in Schlosser and Oerter (2002) and Fernandoy et al. (2010) respectively were merged together. These are referenced in the present study under code c62 (Fig. S3). In this way, the total numbers of δ^{18}O and δ^2H records studied using their 1970–1988 averages were 48 and 21 respectively. Reported dating uncertainty of the set of ice cores was ±1yr in both the Dronning Maud Land (Oerter et al., 2000) and the Ronne Ice Shelf (Graf et al., 1999) for the periods closest to the ones assessed in the present study. Therefore, in the case of the ~20yr averages used in the study, dating uncertainty documented above is expected to be negligible.

It is generally acknowledged that the isotopic composition of meteoric precipitation is related to geographical position (Dansgaard, 1964), and can be statistically modeled employing geographical parameters (Bowen and Revenaugh, 2003). These global trends can indeed be generalized to the Antarctic continent (e.g. Wang et al., 2009). On a regional scale, however, the set of independent variables to describe isotope variations may change. In order to be able to analyze the spatial autocorrelation and derive an isoscape of the stable isotope records, their dependence on geographical factors has to be determined, as in Lorius and Merlivat (1977) or Smith et al. (2002). For the reasons for this and further details, see Section 2.3.

Therefore, in order to determine the geographical factors controlling the ice core water isotopes' variability, latitude (LAT), longitude (LON), elevation (ELE), and distance from the coast (D) were considered in this study. LAT & LON were obtained from the original repository files and converted into meters on a polar stereographic projection with reference to the World Geodetic System 1984 ellipsoid. ELE was extracted from the high-resolution Antarctic digital elevation model of Liu et al. (1999), while D was calculated using the shortest perpendicular distances between the sample points representing the ice cores and the coast line.

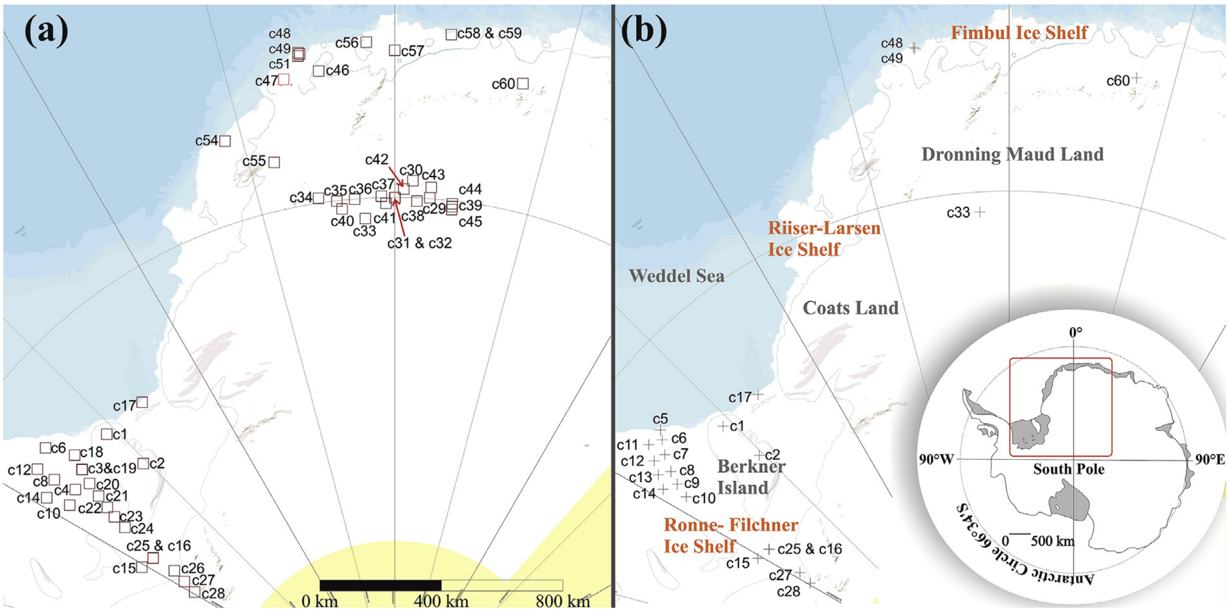


Fig. 1. Spatial distribution of the analyzed (a) δ^{18}O and (b) δ^{2}H ice core records. The study area is marked by a red rectangle on the inset map (bottom right).

2.3. Determination of the geographic factors controlling the water stable isotope variability in firn and ice

In order to obtain representative results on the chosen scale from variography, first the effect of the geographical factors controlling the water stable isotopes' variability in fresh and/or metamorphosed snow has to be minimized. This is because these factors influence the variability of the inspected parameter on a similar and/or larger scale than the phenomena investigated, masking the finer scale pattern, resulting in non-stationarity (Füst and Geiger, 2010; Hohn, 1999).

The following procedures refer only to the δ^{18}O parameter, because after pre-processing, the number of available δ^{2}H records was found to be too low (for details see Section 3.1). In the case of the Antarctic study area, the spatial/geographic dependence of precipitation stable isotope composition is well documented (Lorius and Merlivat, 1977; Masson-Delmotte et al., 2008). In the light of these facts, and following the path indicated by previous studies, multiple regression analysis (Draper and Smith, 1981) was chosen to diminish the influence of topography on such first order factors as e.g. condensation temperature and distillation, leaving the effect of the second order factors such as local air mass trajectories and different moisture sources in the residual field.

However, unlike previous studies, multiple factors (e.g. variance inflation, mean absolute error) were taken into account together - as suggested by O'Brien (2007) - to find the best combination of driving parameters.

2.4. Variography

2.4.1. Theoretical background of the semivariogram

The basic function of geostatistics, the variogram, is a tool for describing the spatial autocorrelation structure of the explored variable and to obtain the weights necessary to be able to predict the values of the Antarctic ice core derived δ^{18}O annual signal at unsampled locations using kriging techniques (Herzfeld, 2004). The variogram can be described mathematically as follows (Molnár et al., 2010): Let $Z(x)$ and $Z(x+h)$ be the values of a parameter sampled at a planar distance $|h|$ from each other. If samples are taken from the same population (stationarity), and they are in accordance with the intrinsic hypothesis of geostatistics, then the variance (VAR) of the difference of $Z(x)$ and $Z(x+h)$ in a given direction is:

$$\begin{aligned}\mathrm{VAR}\,[Z(x+h)-Z(x)] &= \mathrm{VAR}\,[Z(x+h)]+\mathrm{VAR}\,[Z(x)] \\ &\quad -2\mathrm{COV}\,[Z(x+h),\,Z(x)] \\ &= 2\gamma(h)\end{aligned} \tag{1}$$

The function $2\times\gamma(h)$ is called the parameter's variogram, while $\gamma(h)$ is its semivariogram and COV stands for covariance. The semivariogram may be calculated by the Matheron algorithm (Hohn, 1999; Matheron, 1965):

$$\gamma(h) = \frac{1}{2N(h)}\sum_{i=1}^{N(h)}\left[Z(x_i)-Z(x_i+h)\right]^2 \tag{2}$$

where $N(h)$ is the number of lag-h differences, i.e. $n\times(n\text{-}1)/2$ and n corresponds to the number of sites. The most important properties of the function (Fig. 2) are: the value C_0 ("nugget") which withholds information regarding the error of the sampling; the level at which the semivariogram stabilizes is the sill (C: partial sill + C_0: nugget)

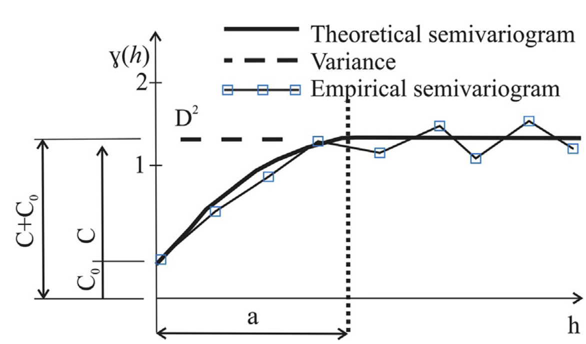


Fig. 2. Properties of the semivariogram, where "a" stands for the range, "C" for the reduced sill and "C_0" for the nugget, if $C_0 > 0$, "h" for lag distance, and "D^2" for the variance of the whole investigated data set; based on Füst and Geiger (2010).

which is equal to the variance for stationary processes, and the range (a) is the distance within which the samples have an influence on each other (Webster and Oliver, 2008) and outside of which they are quasi-independent (Chilès and Delfiner, 2012). This distance (range) determines the average area of influence surrounding the sample locations, within which the measured values of the variable explored are interconnected. In the case of isotropy (Chilès and Delfiner, 2012), the spatial range equals the radius of the area of influence.

If $\gamma(h)$ is a monotonically increasing function (if $h \to \infty$ then $\gamma(h) \to \infty$), the parameter is non-stationary (e.g. in Fig. 3). Moreover, if the semivariogram does not have a rising part, the empirical semivariogram's points will align parallel to the abscissa, giving a nugget-effect type of semivariogram. In this case, the sampling frequency is insufficient to estimate the range (Hatvani et al., 2014).

Empirical semivariograms by themselves are not yet applicable in spatial modeling. They have to be approximated by theoretical functions in order to provide the necessary weights to be used in kriging (Cressie, 1990) for predicting values at unsampled locations (Chilès and Delfiner, 2012; Herzfeld, 2004). However, a thorough discussion of this question is beyond the scope of the present paper.

From the technical perspective, the variogram analysis was conducted on the residuals of the best multiple regression model of the stable isotope records, with a maximum lag distance set to 600 km and 11 uniform bins about 55 km wide. For further details, please see section 2.3.

2.4.2. *Preliminary variography on raw data before minimization of the effect of the geographical factors*

Empirical semivariograms were derived from the original/raw data for δ^{18}O and δ^{2}H. Increasing values of γ were observed for both parameters. For δ^{18}O a strictly monotonic pattern; while for δ^{2}H, two increasing sections were seen: from the smallest lag distance to ~250 km, then from ~370 km onwards (Fig. 3). It should be noted that in the case of δ^{18}O no peaks can be seen as a result of the overwhelming masking effect of geographical factors on water stable isotope variability. Such a semivariogram cannot be used for further evaluation, as explained in Section 2.4.1. Moreover, in the case of δ^{2}H, because of the low number of ice cores (Fig. 1b), γ values could have been calculated for only a few pairs at almost all lag distances (Fig. 3b). Thus, δ^{2}H had to be left out of further analyses.

As discussed in Section 2.3. Geographic factors controlling the water stable isotope variability in firn and ice, and their effect has to be minimized. The previous observations on the particular dataset at hand, therefore, further verify the necessity of the minimization of the determining effect of geographic factors controlling the water stable isotope variability in firn and ice before variography can be commenced.

2.5. *Isoscape derivation*

The procedure of isoscape derivation for the studied Antarctic macro region is based on the methodology used for the global isoscape (Bowen and Wilkinson, 2002) and for an isoscape of an Alpine domain (Kern et al., 2014). The main idea is to:

(i) create an *initial grid* of the stable isotope variance in the region described by the multiple regression model of the supposedly driving geographic variables (LAT, LON, ELE, D). This step was carried out with the ArcGIS Spatial Analyst Raster Calculator tool;
(ii) create an interpolated (ordinary point kriging) *residual grid* using the theoretical semivariogram (Section 2.4) fitted on to the residuals of the multivariate regression model; and
(iii) summarize the corresponding initial (i) and residual (ii) grids to obtain the final map.

Both initial and residual grids were generated uniformly at a resolution of 5 km to facilitate grid calculation. All computations were performed using Golden Software Surfer 11, ArcGIS 10, IBM SPSS 20 and GS+ 10. For certain visualizations of the results, a CorelDRAW Graphics Suite X6 and MS Office 2016 were used.

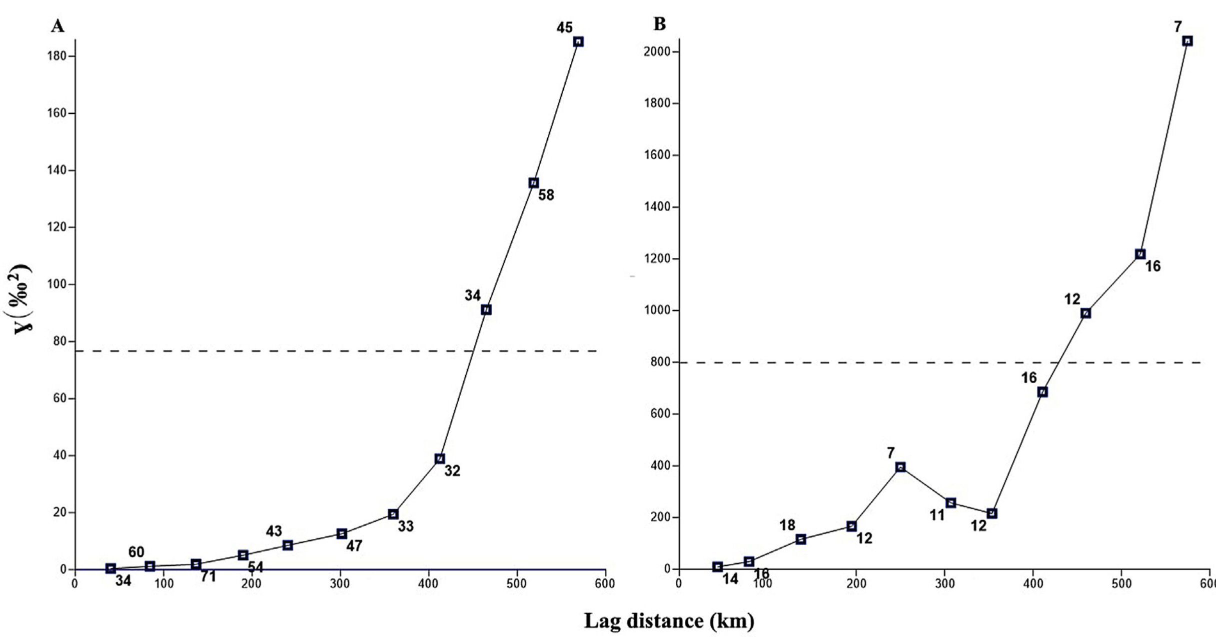


Fig. 3. Empirical semivariogram (squares) derived from the mean original (a) δ^{18}O (averaging period 1970–1988) and (b) δ^{2}H averaging period (1970–1989) values, with the broken line indicating the variance. The numbers indicate the number of data pairs that were used to derive the semivariogram value for a particular bin; bin widths were 55 km.

3. Results

3.1. Minimization of the effect of geographical factors on water stable isotope variability

Motivated by the variogram results on the raw data and the pioneering works of Lorius and Merlivat (1977) and Masson-Delmotte et al. (2008), the geographical factors controlling $\delta^{18}O$ variability in the region were determined/modeled. The values of the multivariate geographical models were subtracted from the averages of firn/ice $\delta^{18}O$ records (raw data). In this way the effects of geographical factors on the averages of firn/ice $\delta^{18}O$ records for the period 1970–1988 were corrected.

Multivariate regression models were computed and compared using independent variables in various combinations of LAT, LON, ELE and D (SOM Table S2). For example, if LON was omitted, adjusted R^2 (R^2_{adj}) = 0.98; mean absolute error (MAE) equals 0.87, and a higher degree of multicollinearity was observed than in the case when LAT was omitted. In fact, the latter case was found to be the most robust choice ($p < 0.01$) for estimating oxygen isotope variation on geographical parameters, $\delta^{18}\widehat{O}$, (Eq. (3)) with an $R^2_{adj} = 0.98$, MAE of 0.95, and an acceptable degree of multicollinearity, which needs to be evaluated in the context of several other factors influencing it (O'Brien, 2007).

$$\delta^{18}\widehat{O} = -(20.51 \pm 0.57) + \left(2.64\,[\pm 0.57] \times 10^{-6}\right) \times LON - \left(5\,[\pm 0.27] \times 10^{-3}\right) \times ELE - \left(1.9\,[\pm 0.12] \times 10^{-5}\right) \times D \tag{3}$$

It should be noted that the uncertainty of the coefficients in the squared brackets is the standard error (SE).

The $R^2_{adj} = 0.98$ may imply that only 2% variance remains in the residuals. However, this is just a method specific and insufficient estimate of the real unexplained spatial variance which is definitely larger (Cressie, 1993). Thus, it has to be explored using variography. The classical non-spatial model (multivariate regression in the present paper) is a special, simplified case of a geostatistical model which is more general (Cressie, 1993). The statistical range of the residuals of the multiple regression spans ~5‰, in addition, their map indicate a spatial structure (Fig. 4a). For instance, negative residuals are observed west of the Berkner Island, or close to zero in Dronning Maud Land and positive ones are clustered south of the Ronne Ice Shelf. The two ice cores on Berkner Island (c1 & c2) gave two of the most positive residuals. This may be explained by the elevated location of ice cores c1 & c2. It suggests that the regional isotopic altitude effect might be unsatisfactory (too steep) here. The regional isotopic altitude effect is mainly determined by the less depleted $\delta^{18}O$ compositions, characterized by the low elevated ice-shelf sites and the more depleted compositions characterized by high elevated central Dronning Maud Land. The explanation for these two positive extremes can be that the hills of the Berkner Island are located right at the edge of the ice shelf, but at a relatively higher latitude. The discrepancy suggest that the main physical parameters (arrival temperature, remaining vapor fraction etc.) have a somewhat different effect on the isotopic Rayleigh process compared to the regional average. The strong local influence clearly lead to deviations and the multiple regression model was unable to follow this microregional pattern (Fig. 4a). Thus, ice cores c1 and c2 were left out during variography, but were included in the spatial interpolation step.

3.2. Variography

The empirical semivariogram of the $\delta^{18}O$ residual was computed with a maximum lag distance set at 600 km, chosen in accordance with the spatial distribution of the cores. The number of cores between 600 and 650 km clearly drops (Fig. S4); as a result, the number of pairs forming the basis of the semivariogram decreases as well at over ~600 km (Fig. 4b). In addition, with the 55 km bins, by keeping the number of pairs relatively even (Fig. 4b) its reliability was ensured. After the empirical semivariogram was obtained, a best-fit spherical model was determined ($R^2 = 0.72$; residual sum of squares was 0.68) following the protocol strongly recommended by Oliver and Webster (2014).

It is clear that the theoretical variogram is not of the nugget-effect type (for a description, see Section 2.4). After a rising part, it stabilizes at a point just slightly above the variance after reaching the sill (Co + C), yielding a roughly 350 km spatial range for the average of the 19 years of $\delta^{18}O$ data used (Fig. 4b). This variogram was later on used to provide the weights for the residual grid derived with ordinary point kriging (Fig. 4c). The standard deviation of the kriging ranged from 0.45 to 1.48 (Fig. 4d). The fact that the kriged map of the residuals (Fig. 4c) reflected a spatial structure and not random noise, clearly indicates that the multivariate regression model was not able to capture this meaningful portion of the spatial variance structure of the firn/ice $\delta^{18}O$.

3.3. Isoscape derivation for $\delta^{18}O$

The initial grid was derived employing LON, ELE and D as the main geographical factors driving the $\delta^{18}O$ variability of firn/ice (Eq (3)). The residual grid was modeled using the weights provided by the variogram of the residual $\delta^{18}O$ (Section 3.2). Afterwards, the initial and residual grids were summed and the regional isoscape of the mean firn/ice $\delta^{18}O$ was obtained (1970–1988).

The values of the residual grid varied by up to 11.8% of the initial grid (Fig. S5), containing a fair portion of the spatial variance of the snow/firn $\delta^{18}O$. Thus, if the residual grid had not been taken into account, this would have been lost. With the average spatial range, the area of influence can then be plotted and unified for the set of examined ice cores. This union indicates the areas where the spatial model is reliable from the geostatistical point of view (Fig. 5), while the map of the standard error of kriging offers an alternative measure of the accuracy of the derived $\delta^{18}O$ isoscape (Fig. 4c).

4. Discussion

4.1. Dependence on geographical factors

Since the main aim of the present study was to determine the spatial range using variography, the effect of the regional geographical factors on firn/ice $\delta^{18}O$ variability had to be minimized. For decades, scientists have been keenly studying the relationship between geographical factors and the stable water isotope composition of surface snow and firn in Antarctica both on a continental and a regional scale, for example Lorius and Merlivat (1977) and the others as may be seen in Table 1.

In regional studies, the number of independent variables applied has in general been smaller than in the present case. In certain Antarctic macro regions the set of geographical factors governing the distribution of stable water isotopes has mainly been determined by elevation (Altnau et al., 2015; Smith et al., 2002). It should be emphasized that in the present case LON, ELE & D explained 98% of the variance in the area explored, just as in the

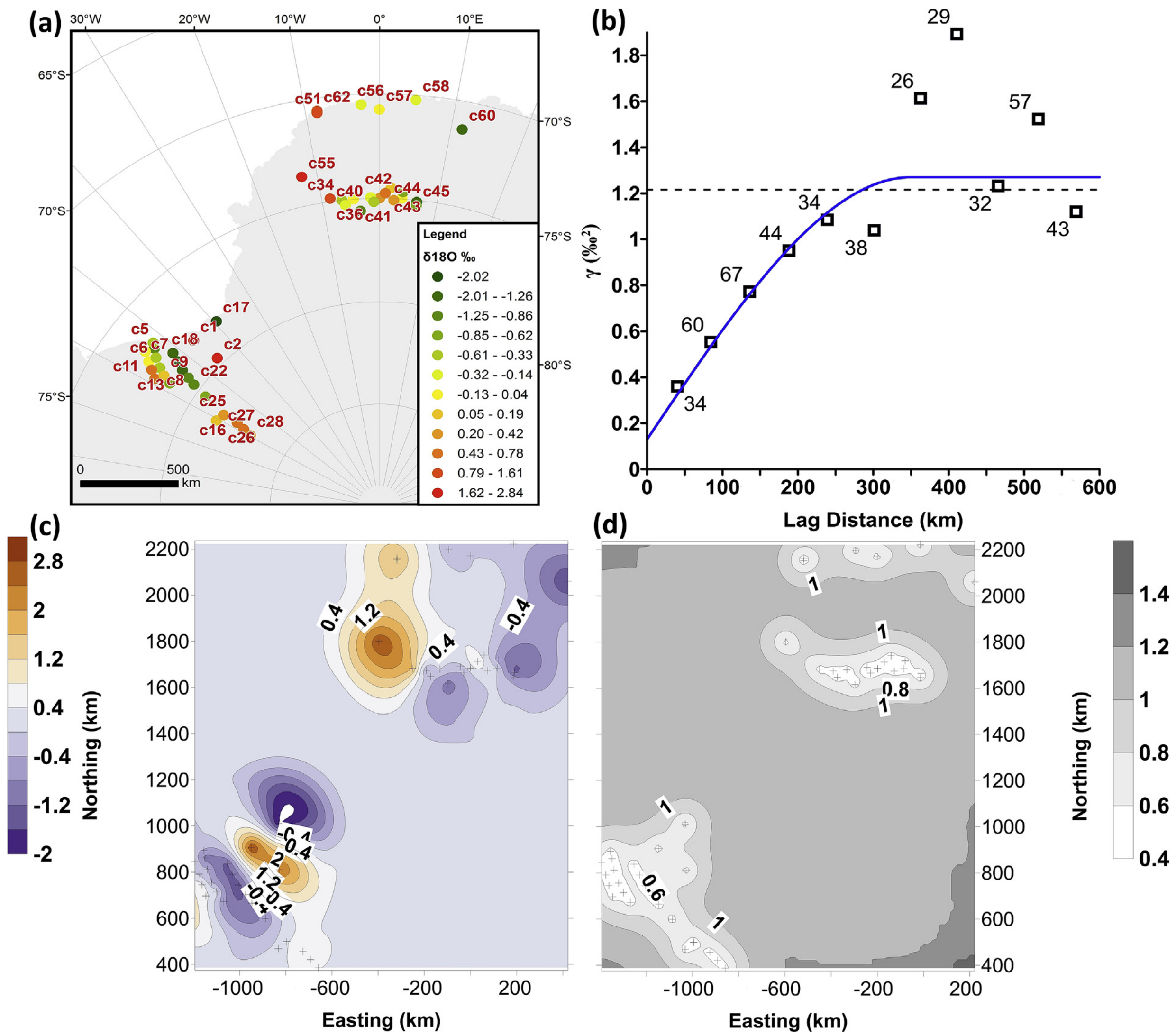


Fig. 4. Map of point residuals, variogram plot, and krigged map of the time average (1970–1988) $\delta^{18}O$ residuals and the standard deviation of kriging. (a) map of point residuals of the multiple regression; (b) empirical (blue squares) semivariogram and spherical theoretical model (blue line) ($C_0 = 0.13$; $C_0 + C = 1.27$; a = 350 km; $r^2 = 0.72$; bin width ~55 km) of the residuals. The variance is marked by the broken line. The numbers indicate the number of data pairs that were used to derive the variogram value (blue squares) for a particular bin in (b). (c) The ordinary point kriged map of the residuals; (d) the standard deviation of kriging. The faded crosses mark the locations of the ice cores in (c) & (d). Easting is the distance from the Prime Meridian, negative towards W and positive to E, while Northing is the distance from the South Pole, both in km. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

work of Altnau et al. (2015), which overlapped with a fair portion of the western part of our studied region.

In general the explanatory power of the models in these studies was lower than those used here, and although there were cases when multiple criteria were used to evaluate the models (Wang et al., 2010), this was not the most common practice. Usually R^2 was used (e.g. Altnau et al. (2015); Masson-Delmotte et al. (2008)), and despite its importance (O'Brien, 2007), no attention was paid to multicollinearity, unlike in the present study. In addition, the data gathered here were homogeneous, inasmuch as firn/ice core data uniformly averaged for 1970–1988 were used. Casual snow samples or averages of hundreds of meters of ice cores were omitted to avoid potential bias towards a short or extraordinarily long time frame. These considerations make the presented approach more reliable for the region than any other previously.

A comparison with previous studies partially overlapping in time and/or space on the dependence of these geographical factors led to meaningful results only in the case of elevation. The value of 0.005‰ m^{-1} obtained in the present study was of the same order as those in the literature, but was one of the lowest among them even if uncertainty (SE in Eq. (3)) is taken into consideration. In particular, for the investigation of areas ranging between the coastline and the East Antarctic Plateau (Altnau et al., 2015; Smith et al., 2002), the coefficient of elevation (0.008‰ m^{-1}) was higher than that in the present study. For research conducted on a continental scale, the derived coefficients have been lower (0.007‰ m^{-1} Masson-Delmotte et al. (2008) and 0.0068‰ m^{-1} (Wang et al., 2010)). Our lower value can be explained by the fact that only a small part of the area investigated here extends onto the East Antarctic Plateau where a steeper elevation effect was observed.

The two most outlying model residuals suggested that the multivariate regression applied here was unable to accurately capture the microregional characteristics of the elevated relief of Berkner Island (c1 & c2) rising between the Ronne and Filchner Ice Shelves (Fig. 1). These regional differences lead to the omission of ice cores c1 and c2 from the multiple regression model.

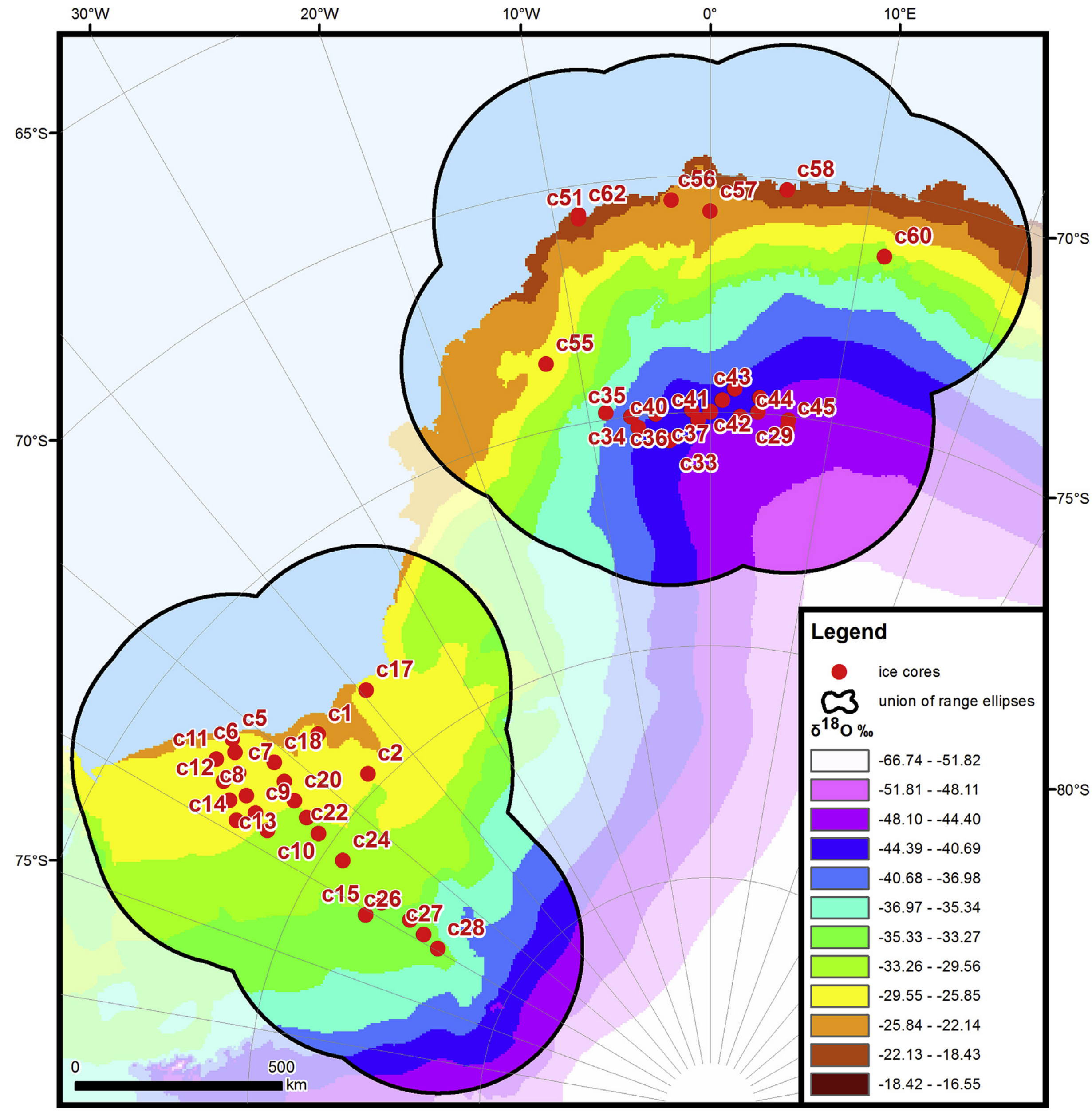


Fig. 5. Isoscape of $\delta^{18}O$ for the region for 1970–1988. The union of the 350 km range ellipses is marked with a black line. Isotropy was assumed, as the directional subsets proved insufficient in the course of the exploration of anisotropy. Red dots mark the ice cores used in the study. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Regression models of water stable isotopes using geographical variables in Antarctica overlapping with the study area.

Study	Used independent variables	Estimated dependent variables	Scale
Masson-Delmotte et al., 2008	sin (LAT), ELE, D	$\delta^{18}O$, $\delta^{2}H$	Continental
Wang et al., 2010	LAT, LON, ELE, D	$\delta^{18}O$, $\delta^{2}H$	
Wang et al, 2009	$\lvert\sin(LAT)\rvert^2$, $\lvert\sin(LAT)\rvert$, ELE	$\delta^{18}O$	
Altnau et al., 2015	ELE	$\delta^{18}O$	Regional
This study	LON, ELE, D	$\delta^{18}O$	

On the basis of the results presented here, it may be suspected that on the scale of the investigated region LON, distance from the coast (D) and elevation (ELE) are accounted as determining most of the variance of $\delta^{18}O$ attributable to geographical factors in the area. This makes sense, since, from a physical point of view, elevation is a main determinant of condensation temperature. The distance from coast controls moisture loss efficiency related to sequential precipitation events on the course of the inland transport of the air parcel. However, these factors do not fully account for second-order controls e.g. specific local air mass trajectories, the differing seasonality of moisture sources for precipitation in Antarctica (Sodemann and Stohl, 2009). By removing the driving effect of the geographical (first order) factors with Eq. (3), the net effect of the previously mentioned second order factors (see section 2.3) was retained in the residuals. These ultimately provided the input values for the variography and consequently the residual grid when the isoscape was modeled.

4.2. Variography and isoscape derivation

4.2.1. Variography and kriging

After the governing effect of geographical factors on firn/ice δ^{18}O variability had been minimized, the dataset was prepared for variography. In contrast to mathematical interpolation, where the same algorithm is applied to every location, geostatistical interpolation using variograms is able to take regional properties into account (Herzfeld, 2004). It provides a spatially more data adaptive approach as underlined by Wang et al. (2010) in his research on water stable isotopes.

The statistical uncertainty of the interpolation may come from variography: the nugget ($\sqrt{C_0} = 0.36$‰) referring to the sampling error and the kriging standard deviation (KSD; higher than 0.45‰; Fig. 4c) are governed by the spatial distribution of the ice cores (Chilès and Delfiner, 2012). In an ideal case, exact interpolation is indicated by KSD being zero (Wackernagel, 2003), which is anyhow unlikely in nature. If $C_0 > 0$, as in the present case (Fig. 4a), this was impossible. These uncertainties were acceptable, since the usual analytical accuracy of the stable oxygen isotope analysis is ~0.1–0.2‰ depending on the applied isotope analytical method (Lis et al., 2008). In addition, the well-known stratigraphic noise as a natural factor in ice core records (Fisher et al., 1985) also affected the firn δ^{18}O, causing signal disturbance over small distances.

4.2.2. On isoscape and spatial range

The presented δ^{18}O isoscape describes an Antarctic macro region, 20% of which (~0.5 × 10^6 km^2) lies on ice shelves (Scambos et al., 2007). Oddly, it is not covered by the continental δ^{18}O maps (Wang et al., 2009, 2010), despite the existence of data. As expected, the mean stable isotope content indicated by the isoscape (Fig. 5) decreased with increasing elevation & distance, and with temperatures decreasing inland. These phenomena can be explained by isotopic fractionation inducing the preferential condensation of heavier isotopologues during precipitation processes, leading to depleted water stable isotope composition for both water vapor and precipitations penetrating the continent.

Ambient temperature is a primary physical factor in determining δ^{18}O variability in precipitation and therefore also influences surface snow and consequently firn/ice isotope variability. Thus, it was interesting to compare the spatial characteristics of δ^{18}O in surface snow and temperature. Unfortunately, due to the lack of comparable direct instrumental temperature measurements in space and time for such purposes, monthly mean near surface (2 m) air temperature, representing the region of interest from 1970 to 1988 had to be used. This data was extracted from NCEP/NCAR Reanalysis Products (Kalnay et al., 1996). The same data treatment was applied for the temperature field as for the δ^{18}O records of the ice cores, and variography was conducted on the residuals of its multiple regression model (Table S3). The spatial range of the temperatures at the near surface- (2 m) was 428 km (Fig. 6).

It is well-known that the inversion and surface temperature are well correlated in Antarctica (Jouzel et al., 1987), exhibiting a gradient of 0.67 for correspondent temperature changes with changes at the surface being larger. Therefore, a positive correlation between water isotope composition and surface temperature is expected.

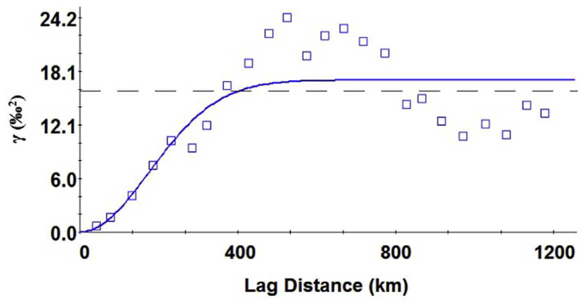


Fig. 6. Empirical semivariogram of residual near surface (2 m) temperature data (see Table S3) for the region (1970–1988) marked with black squares and the fitted theoretical Gaussian model (blue line).

A recent field study (Steen-Larsen et al., 2014) documented in the upper most 0.5 cm of snow an imprint of the isotopic composition of vapor, by vapor exchange even between precipitation events pointing to the imprint of ambient temperature variations via fractionation processes. This surface processes can modify the temperature signal imprinted into the stable water isotopes transported to the surface by precipitation from the cloud formation level. This is further supported by the statistical evidence of related data (Hatvani and Kern, 2017). The similar spatial variability domains for firn/ice δ^{18}O (350 km) and for the near surface temperature (428 km) might be an indication that most of the observed water isotope variations are temperature driven, yet the shorter domain for δ^{18}O calls for additional factors to be involved.

Indeed, other phenomena, such as post deposition processes (e.g. stratigraphic noise (Fisher et al., 1985)), can be also considered as major factors in the determination of δ^{18}O variability. It is interesting to note that a variogram analysis of surface snow accumulation in an area overlapping with the domain of the present study reported the effective radii of spatial autocorrelation to be 200–250 km from NE to SW and 100–150 km perpendicular to this direction (Rotschky et al., 2007). The previously mentioned stratigraphic noise is therefore a major factor in driving the spatial variability of snow accumulation.

Thus, the intermediate "position" of the firn/ice δ^{18}O spatial range between the ranges characterizing the spatial variability of air temperature and snow accumulation reinforces the notion that a combination of these processes is responsible for firn/ice δ^{18}O spatial variability.

4.2.3. A practical message for future site selection

If the aim is to study spatial (from regional to continental) variations of ice/firn stable isotopes one or two drilling sites are unsatisfactory. Therefore, in the case of an array of ice cores gesostatistical planning is inevitable to optimize sampling, as done in the case of the International Trans-Antarctic Scientific Expedition (Cressie, 1998). In the presented region of Antarctica, the 350 km spatial range conveys the clear message that the current set of ice cores does not cover the whole sector. The area outside the coverage of the cores (outside the union of the range ellipses (Fig. 5) is where new ice cores can contribute the most to the ice core network as an additional geostatistical criterion for optimal site selection (Vance et al., 2016). The drilling of additional cores would be helpful in the improvement of the description of spatial variability, as has also been suggested by the IPICS (2006) initiative. If samples are taken by interpolation in the areas outside the range ellipses' union, the interdependence structure of firn/ice δ^{18}O variability remains undetectable, as samples there are already quasi-independent (Chilès and Delfiner, 2012). Naturally, if new cores are drilled, it may be presumed that the spatial autocorrelation structure of the dataset will change, and therefore after any such campaign, a recalibration would be required.

5. Conclusions

Using the obtained dataset, the regional dependence of firn/ice $\delta^{18}O$ variability on geographical factors was determined. It was shown that in the study area $\delta^{18}O$ variance was most precisely estimated - taking multicollinearity into account as well - by employing longitude, elevation and the distance from the coast in a multivariate regression model. Consequently, after correction for the effect of the geographic influence, the spatial autocorrelation structure was revealed using variography, serving as the basis for isoscape derivation employing kriging. The 350 km spatial range explicitly indicates the areas not as yet covered by the currently available network of ice cores. Furthermore, the geostatistical findings support the notion that the combination of the spatial variability of air temperature and snow accumulation is likely to regulate firn/ice $\delta^{18}O$ spatial variability. These results bring us closer to the accomplishment of one of the ultimate aims of research in this field, an improved continental isoscape for Antarctica, since it has so far been neglected in global isoscapes (Terzer et al., 2013; van der Veer et al., 2009).

Furthermore, we provide additional information about the spatial extent of a common water stable isotope signal in Antarctica and offer guidance to experimenters on where to drill additional cores in order to improve the overall representativeness of the shallow Antarctic ice/firn core network. Moreover, the derived isoscape can be used as a benchmark in model validation of isotope enabled GCMs in the case of scarce station based precipitation stable isotope records over Antarctica.

Acknowledgements

We the authors would like to thank Paul Thatcher for his work on our English version and say thanks for the stimulating discussion with Julien P. Nicolas from the Ohio State University. The work of I.G. Hatvani was supported within the framework of TÁMOP 4.2.4. A/1-11-1-2012-0001 „National Excellence Program – Elaborating and operating an inland student and researcher personal support system". The project was subsidized by the European Union and co-financed by the European Social Fund. Special thanks to D. Divine, C.M. Laluraj for providing access to their data and the researchers who archived their Antarctic ice core derived stable isotope data in PANGAEA. The NCEP Reanalysis data was provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their website at http://www.esrl.noaa.gov/psd/. We would also like to give thanks for the support of the MTA "Lendület" program (LP2012-27/2012) and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. This is contribution No. 39 of 2ka Palæoclimate Research Group.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.polar.2017.04.001.

References

Altnau, S., Schlosser, E., Isaksson, E., Divine, D., 2015. Climatic signals from 76 shallow firn cores in dronning Maud Land, East Antarctica. Cryosphere 9, 925–944.
Araguás-Araguás, L., Froehlich, K., Rozanski, K., 2000. Deuterium and oxygen-18 isotope composition of precipitation and atmospheric moisture. Hydrol. Process. 14, 1341–1355.
Bowen, G.J., 2010. Isoscapes: spatial pattern in isotopic biogeochemistry. Annu. Rev. Earth Planet. Sci. 38, 161–187.
Bowen, G.J., Revenaugh, J., 2003. Interpolating the isotopic composition of modern meteoric precipitation. Water Resour. Res. 39, 1299.
Bowen, G.J., Wilkinson, B., 2002. Spatial distribution of δ18O in meteoric precipitation. Geology 30, 315–318.
Chilès, J.-P., Delfiner, P., 2012. Geostatistics. Wiley, Canada.
Coplen, T.B., 1994. Reporting of stable hydrogen, carbon and oxygen isotopic abundances. Pure App. Chem. 66, 273–276.
Cressie, N., 1990. The origins of kriging. Math. Geol. 22, 239–252.
Cressie, N., 1993. Statistics for Spatial Data. John Wiley & Sons, Inc.
Cressie, N., 1998. Transect-spacing design of ice cores on the antarctic continent. Can. J. Statistics 26, 405–418.
Dansgaard, W., 1964. Stable isotopes in precipitation. Tellus 16, 436–468.
Divine, D.V., Isaksson, E., Kaczmarska, M., Godtliebsen, F., Oerter, H., Schlosser, E., Johnsen, S.J., van den Broeke, M., van de Wal, R.S.W., 2009. Tropical pacific - high latitude SouthAtlantic teleconnections as seen in the δ18O variability in antarctic coastal ice cores. J. Geophys.Res 114, D11112.
Draper, N.R., Smith, H., 1981. Applied Regression Analysis. Wiley.
EPICA, 2004. Eight glacial cycles from an Antarctic ice core. Nature 429, 623–628.
EPICA, 2006. One-to-one coupling of glacial climate variability in Greenland and Antarctica. Nature 444, 195–198.
Fernandoy, F., Meyer, H., Oerter, H., Wilhelms, F., Graf, W., Schwander, J., 2010. Temporal and spatial variation of stable-isotope ratios and accumulation rates in the hinterland of Neumayer station, East Antarctica. J. Glaciol. 56, 673–687.
Fisher, D.A., Reeh, N., Clausen, H.B., 1985. Stratigraphic noise in time series derived from ice cores. Ann. Glaciol. 7, 76–83.
Füst, A., Geiger, J., 2010. Monitoring planning and evaluation using geostatistics, I. Geostatistical support for verification sampling based on professional opinion. Földtani Közlöny 140, 303–312.
Graf, W., Moser, H., Reinwarth, O., Kipfstuhl, J., Oerter, H., Minikin, A., Wagenbach, D., 1994. Snow-accumulation rates and isotopic content (2H, 3H) of near-surface firn from the Filchner-Ronne Ice Shelf, Antarctica. Ann. Glaciol. 20, 121–128.
Graf, W., Reinwarth, O., Oerter, H., Mayer, C., Lambrecht, A., 1999. Surface accumulation on foundation ice stream, Antarctica. Ann. Glaciol. 29, 23–28.
Hammer, C.U., 1989. Dating by physical and chemical seasonal variations and reference horizons. In: Oeschger, H, Langway, C.C. (Eds.), The Environmental Record in Glaciers and Ice Sheets. John Wiley & Sons Limited, pp. 99–121.
Hatvani, I.G., Kern, Z., 2017. Weighting alternatives for water stable isotopes: statistical comparison between station- and firn/ice records. Pol. Polar Res. 38 (2) http://dx.doi.org/10.1515/popore-2017-0006 (in press).
Hatvani, I.G., Magyar, N., Zessner, M., Kovács, J., Blaschke, A.P., 2014. The Water Framework Directive: can more information be extracted from groundwater data? A case study of Seewinkel, Burgenland, eastern Austria. Hydrogeol. J. 22, 779–794.
Herzfeld, U.C., 2004. Atlas of Antarctica: Topographic Maps from Geostatistical Analysis of Satellite Radar Altimeter Data : with 169 Figures. Springer.
Hohn, M.E., 1999. Geostatistics and Petroleum Geology, 2ed. Springer Science+-Business Media Dordrecht, The Netherlands.
IPICS, 2006. The IPICS 2k Array: a Network of Ice Core Climate and Climate Forcing Records for the Last Two Millennia. International Partnerships in Ice Coring Sciences, p. 4.
Jones, P.D., Briffa, K.R., Osborn, T.J., Lough, J.M., van Ommen, T.D., Vinther, B.M., Luterbacher, J., Wahl, E.R., Zwiers, F.W., Mann, M.E., Schmidt, G.A., Ammann, C.M., Buckley, B.M., Cobb, K.M., Esper, J., Goosse, H., Graham, N., Jansen, E., Kiefer, T., Kull, C., Küttel, M., Mosley-Thompson, E., Overpeck, J.T., Riedwyl, N., Schulz, M., Tudhope, A.W., Villalba, R., Wanner, H., Wolff, E., Xoplaki, E., 2009. High-resolution palaeoclimatology of the last millennium: a review of current status and future prospects. Holocene 19, 3–49.
Jouzel, J., 2013. A brief history of ice core science over the last 50 yr. Clim. Past. 9, 2525–2547.
Jouzel, J., Lorius, C., Petit, J.R., Genthon, C., Barkov, N.I., Kotlyakov, V.M., Petrov, V.M., 1987. Vostok ice core: a continuous isotope temperature record over the last climatic cycle (160,000 years). Nature 329, 402–408.
Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Jenne, R., Joseph, D., 1996. The NCEP/NCAR 40-year Reanalysis project. Bull. Am. Meteorological Soc. 77, 437–471.
Kern, Z., Kohán, B., Leuenberger, M., 2014. Precipitation isoscape of high reliefs: interpolation scheme designed and tested for monthly resolved precipitation oxygen isotope records of an Alpine domain. Atmos. Chem. Phys. 14, 1897–1907.
Kuramoto, T., Goto-Azuma, K., Hirabayashi, M., Miyake, T.i., Motoyama, H., Dahl-Jensen, D., Steffensen, J.P., 2011. Seasonal variations of snow chemistry at NEEM. Greenl. Ann. Glaciol. 52, 193–200.
Lachniet, M.S., Lawson, D.E., Stephen, H., Sloat, A.R., Patterson, W.P., 2016. Isoscapes of δ18O and δ2H reveal climatic forcings on Alaska and Yukon precipitation. Water Resour. Res. 52, 6575–6586.
Lis, G., Wassenaar, L.I., Hendry, M.J., 2008. High-Precision laser spectroscopy D/H and 18O/16O measurements of microliter natural water samples. Anal. Chem. 80, 287–293.
Liu, H., Jezek, K.C., Li, B., 1999. Development of an Antarctic digital elevation model by integrating cartographic and remotely sensed data: a geographic information system based approach. J. Geophys. Res: Solid earth. 104, 23199–23213.
Lorius, C., Merlivat, L., 1977. Distribution of mean surface stable isotope values in

East Antarctica: observed changes with depth in the coastal area. Int. Assoc. Hydrological Sci. Publ. 118, 127–137.
Masson-Delmotte, V., Hou, S., Ekaykin, A., Jouzel, J., Aristarain, A., Bernardo, R.T., Bromwich, D., Cattani, O., Delmotte, M., Falourd, S., Frezzotti, M., Gallée, H., Genoni, L., Isaksson, E., Landais, A., Helsen, M.M., Hoffmann, G., Lopez, J., Morgan, V., Motoyama, H., Noone, D., Oerter, H., Petit, J.R., Royer, A., Uemura, R., Schmidt, G.A., Schlosser, E., Simões, J.C., Steig, E.J., Stenni, B., Stievenard, M., van den Broeke, M.R., van de Wal, R.S.W., van de Berg, W.J., Vimeux, F., White, J.W.C., 2008. A review of Antarctic surface snow isotopic composition: observations, atmospheric circulation, and isotopic modeling. J. Clim. 21, 3359–3387.
Matheron, G., 1965. Les Variables régionalisées et leur estimation: une application de la théorie des fonctions aléatoires aux sciences de la nature. Masson et Cie.
Merlivat, L., Jouzel, J., 1979. Global climatic interpretation of the deuterium-oxygen 18 relationship for precipitation. J. Geophys. Res. Oceans 84, 5029–5033.
Molnár, S., Füst, A., Szidarovszky, F., Molnár, M., 2010. Models in Environmental Informatics II. Szent István University, Department of Informatics, Göldöllő, Hungary.
Naik, S.S., Thamban, M., Laluraj, C.M., Redkar, B.L., Chaturvedi, A., 2010. A century of climate variability in central dronning Maud Land, East Antarctica, and its relation to southern annular mode and el niño-southern oscillation. J. Geophys. Res: Atmos. 115, D16102.
NGRIP, 2004. High-resolution record of Northern Hemisphere climate extending into the last interglacial period. Nature 431, 147–151.
NOAA, 2014. Paleoclimatology Database. www.ncdc.noaa.gov/paleo/icecore/current.html. Accessed 01 February 2014.
O'Brien, R.M., 2007. A caution regarding rules of thumb for variance inflation factors. Qual. Quantity 41, 673–690.
Oerter, H., Wilhelms, F., Jung-Rothenhäusler, F., Göktas, F., Miller, H., Graf, W., Sommer, S., 2000. Accumulation rates in Dronning Maud Land, Antarctica, as revealed by dielectric-profiling measurements of shallow firn cores. Ann. Glaciol. 30, 27–34.
Oliver, M.A., Webster, R., 2014. A tutorial guide to geostatistics: computing and modelling variograms and kriging. CATENA 113, 56–69.
PANGAEA, 2014. Paleoclimatology Database. http://www.pangaea.de/. Accessed 01 January 2014.
Rotschky, G., Holmlund, P., Isaksson, E., Mulvaney, R., Oerter, H., Van den Broeke, M.R., Winther, J.-G., 2007. A new surface accumulation map for western Dronning Maud Land, Antarctica, from interpolation of point measurements. J. Glaciol. 53, 385–398.
Scambos, T.A., Haran, T.M., Fahnestock, M.A., Painter, T.H., Bohlander, J., 2007. MODIS-based Mosaic of Antarctica (MOA) data sets: continent-wide surface morphology and snow grain size. Remote Sens. Environ. 111, 242–257.
Schlosser, E., Oerter, H., 2002. Shallow firn cones from Neumayer, Ekströmisen, Antartica: a comparison of accumulation rates and stable-isotope ratios. Ann. Glaciol. 35, 91–96.
Sjolte, J., Hoffmann, G., Johnsen, S.J., Vinther, B.M., Masson-Delmotte, V., Sturm, C., 2011. Modeling the water isotopes in Greenland precipitation 1959–2001 with the meso-scale model REMO-iso. J. Geophys. Res. Atmos. 116 (n/a-n/a).
Smith, B.T., Van Ommen, T.D., Morgan, V.I., 2002. Distribution of oxygen isotope ratios and snow accumulation rates in Wilhelm II Land, East Antarctica. Ann. Glaciol. 35, 107–110.
Sodemann, H., Stohl, A., 2009. Asymmetries in the moisture origin of Antarctic precipitation. Geophys. Res. Lett. 36, L22803.
Steen-Larsen, H.C., Masson-Delmotte, V., Hirabayashi, M., Winkler, R., Satow, K., Prié, F., Bayou, N., Brun, E., Cuffey, K.M., Dahl-Jensen, D., Dumont, M., Guillevic, M., Kipfstuhl, S., Landais, A., Popp, T., Risi, C., Steffen, K., Stenni, B., Sveinbjörnsdottír, A.E., 2014. What controls the isotopic composition of Greenland surface snow? Clim. Past 10, 377–392.
Steig, E.J., Mayewski, P.A., Dixon, D.A., Kaspari, S.D., Frey, M.M., Schneider, D.P., Arcone, S.A., Hamilton, G.S., Spikes, V.B., Albert, M., Meese, D.A., Gow, A.J., Shuman, C.A., White, J.W.C., Sneed, S., Flaherty, J., Wumkes, M., 2005. High-resolution ice cores from US ITASE (West Antarctica): development and validation of chronologies and determination of precision and accuracy. Ann. Glaciol. 41, 77–84.
Sturm, K., Hoffmann, G., Langmann, B., Stichler, W., 2005. Simulation of δ18O in precipitation by the regional circulation model REMOiso. Hydrol. Process. 19, 3425–3444.
Terzer, S., Wassenaar, L.I., Araguás-Araguás, L.J., Aggarwal, P.K., 2013. Global isoscapes for δ18O and δ2H in precipitation: improved prediction using regionalized climatic regression models. Hydrol. Earth Syst. Sci. 17, 4713–4728.
van der Veer, G., Voerkelius, S., Lorentz, G., Heiss, G., Hoogewerff, J.A., 2009. Spatial interpolation of the deuterium and oxygen-18 composition of global precipitation using temperature as ancillary variable. J. Geochem. Explor. 101, 175–184.
Vance, T.R., Roberts, J.L., Moy, A.D., Curran, M.A.J., Tozer, C.R., Gallant, A.J.E., Abram, N.J., van Ommen, T.D., Young, D.A., Grima, C., Blankenship, D.D., Siegert, M.J., 2016. Optimal site selection for a high-resolution ice core record in East Antarctica. Clim. Past. 12, 595–610.
Vaughan, D.G., Bamber, J.L., Giovinetto, M., Russell, J., Cooper, A.P.R., 1999. Reassessment of net surface mass balance in Antarctica. J. Clim. 12, 933–946.
Wackernagel, H., 2003. Multivariate Geostatistics, 3 ed. Springer-Verlag Berlin Heidelberg.
Wang, Y., Hou, S., Masson-Delmotte, V., Jouzel, J., 2009. A new spatial distribution map of δ18O in Antarctic surface snow. Geophys. Res. Lett. 36, L06501.
Wang, Y., Hou, S., Masson-Delmotte, V., Jouzel, J., 2010. A generalized additive model for the spatial distribution of stable isotopic composition in Antarctic surface snow. Chem. Geol. 271, 133–141.
Webster, R., Oliver, M.A., 2008. Geostatistics for Environmental Scientists, 2 ed. John Wiley & Sons, Ltd.
Werner, M., Heimann, M., 2002. Modeling interannual variability of water isotopes in Greenland and Antarctica. J. Geophys. Res. Atmos. 107 (D1), 1–13.
Werner, M., Langebroek, P.M., Carlsen, T., Herold, M., Lohmann, G., 2011. Stable water isotopes in the ECHAM5 general circulation model: toward high-resolution isotope modeling on a global scale. J. Geophys. Res. Atmos. 116.
Wolff, E.W., Chappellaz, J., Blunier, T., Rasmussen, S.O., Svensson, A., 2010. Millennial-scale variability during the last glacial: the ice core record. Quat. Sci. Rev. 29, 2828–2838.
Xi, X., 2014. A review of water isotopes in atmospheric general circulation models: recent advances and future prospects. Int. J. Atmos. Sci. 2014, 16.